

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:43:31 ; Search time 41 Seconds  
(without alignments)

940.800 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNHGPDATEAEEDVDFMTV.....VTDEIVXEFMTPRKLSFDQ 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619-526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 78:\*  
1: piri:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	99.3	471	1 A41706	tryptophan-tRNA li
2	2028.5	95.9	475	1 YWBO	tryptophan-tRNA li
3	1946.5	92.0	475	1 YWRBPR	tryptophan-tRNA li
4	1938	91.6	481	2 S50053	tryptophan-tRNA li
5	1210	57.2	395	2 S58157	hypothetical prote
6	1163	55.0	432	2 S51901	tryptophan-tRNA li
7	967	42.9	386	2 C90190	tryptophan-tRNA li
8	803	37.9	395	2 C75020	tryptophanyl-tRNA
9	626.5	29.6	301	2 G71206	tryptophan-tRNA li
10	534.5	25.3	380	2 G84373	tryptophanyl-tRNA
11	409.5	19.4	370	2 F64476	tryptophan-tRNA li
12	397.5	18.8	364	2 B69131	tryptophan-tRNA li
13	386	18.2	134	2 T43806	tryptophan-tRNA li
14	372.5	17.5	420	2 B69461	tryptophanyl-tRNA
15	363	16.7	374	2 D72477	probable tryptopha
16	263.5	12.7	533	2 F84371	tryptophanyl-tRNA
17	192	9.1	323	2 H69346	tyrosyl-tRNA synth
18	188	8.9	364	2 E72512	probable tyrosyl-t
19	185.5	8.8	341	2 D95260	tryptophanyl-tRNA
20	185.5	8.8	341	2 G98125	tryptophan-tRNA li
21	174.5	8.2	341	2 B86633	tryptophan-tRNA li
22	172	8.1	334	2 A45993	tyrosine-tRNA liga
23	170.5	8.1	366	2 S75410	tryptophan-tRNA li
24	169	8.0	335	2 R70385	tryptophan-tRNA li
25	167	7.9	351	2 E75438	tryptophanyl-tRNA
26	166.5	7.9	346	2 B71496	tryptophan-tRNA li
27	164.5	7.8	335	2 A86410	protein F3M18.22 [
28	163.5	7.7	333	2 E70100	tryptophan-tRNA li
29	163	7.7	460	2 C84750	probable tyrosyl-t

30 159.5 7.5 337 2 F71300  
31 158.5 7.5 375 2 B75072  
32 158 7.5 346 2 C81654  
33 154.5 7.3 102 2 T44994  
34 152 7.2 344 2 H86590  
35 152 7.2 344 2 C72034  
36 149.5 7.1 327 2 C84374  
37 149.5 7.1 408 2 T03741  
38 146 6.9 339 2 B64676  
39 146 6.9 375 2 F71093  
40 145 6.9 337 2 A11066  
41 138 6.5 319 2 H69102  
42 135.5 6.4 328 2 C72370  
43 135 6.4 365 2 E82052  
44 132.5 6.3 343 2 S73024  
45 132 6.2 328 1 YWBSF

#### ALIGNMENTS

##### RESULT 1

241706  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homol  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text change 03-Jun-2002  
C:Accession: A41633; A41706; S19246; JNC676; JH0533; S26287  
R:Fleckenner, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11523-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma  
A:Reference number: A41633; MUID:92107982; PMID:1763065  
A:Accession: A41633  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <FILE>  
A:Cross-references: GB:X59892; NID:g33820; PIDN:CAA42545.1; PID:g30821  
R:Rubin, B.V.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophanyl-tRNA synthetase expression in human fibroblasts  
A:Reference number: A41706; XJID:92105071; PMID:1761529  
A:Accession: A41706  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:g184656; PIDN:AAA67324.1; PID:g184657  
R:Buwitt, U.; Flohr, T.; Boettger, E.C.  
EMBO J. 11, 489-496, 1992  
A:Title: Molecular cloning and characterization of an interferon induced human cDNA with  
A:Reference number: S19246; MUID:92164636; PMID:1537332  
A:Accession: S19246  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-423, R, 425-471 <BUW>  
A:Cross-references: EMBL:X65570; NID:g32708; PIDN:CAA44450.1; PID:g32709  
R:Prolova, L.Y.; Grigorova, A.V.; Sudomocina, M.A.; Kisselev, L.L.  
Gene 128, 217-245, 1993  
A:Title: The human gene encoding tryptophanyl-tRNA synthetase: Interferon-response eleme  
A:Reference number: JN0676; MUID:93292992; PMID:7695728  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141,182-471 <PROI>  
A:Cross-references: GB:X67918; GB:S62837; NID:g37968; GB:X67919; NID:g37969; GB:X67920; N  
4; NID:g37974; GB:X67925; GB:S62855; NID:g37975; GB:X67926; GB:S62856; NID:g37976; GB:X6  
A:Note: The authors translated the codon GGG for residue 55 as Cys and GAG for residue 34  
R:Prolova, L.Y.; Sudomocina, M.A.; Grigorova, A.V.; Zinovieva, O.L.; Kisselev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human trypt  
A:Reference number: JH0533; MUID:92112058; PMID:1765274  
A:Accession: JH0533  
A:Molecule type: mRNA



[illegible]

124 DDEKELFKQGVSLDQCFARENAKDIIVAGFDPKKTFFIMNSTYVG--GAFYQNVRLA 181  
186 KHVTFNQVKGIFGFTSDICIGKISPAIAQAASFNSPQIFRDRDIOCLIPCAIDQDP 245  
182 KCITANQSKACGFTSDSISGKIHFAISIAAASFSSFPPIHFNAGAKDIFCLIPCAIDQDP 241  
246 YFRMTDVAIPRIGYPKAPALLHSTFFPALQCAQTKMSADPNSSIFLTDIAKQIKTKVNKH 305  
242 YRLTEDVSGRLKFKPKAPALLHSTFFPALQCAQTKMSADPNSSIFLTDIAKQIKTKVNKH 301  
306 AFSGGRTTIEHRQFGNCDVDSFMYLTFEEDDKLEQIRKDYTGSGAMLTGELKKALI 365  
302 AFSGGGATIEHREKGNPDVAVAYQFLSFFLDDDEKLAQLVNTYKAGTILSTGEMGECI 361  
366 EYLOPLIAHQARRKEVTDEIVKEEVT-PRKLSF 398  
362 KULQFVDFQARSKVDEATLDMFMGGRKLEW 395

RESULT 6  
S51901  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - yeast (*Saccharomyces cerevisiae*)  
N:Altername names: protein HR8432; protein O0992; tryptophanyl-tRNA synthetase  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 05-May-1995 #sequence revision 03-Aug-1995 #text\_change 03-Jun-2002  
C:Accession: S51901; S59177; S66793  
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
submitted to the EMBL Data Library, January 1995  
A:Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including  
and a Delta.  
A:Reference number: S51848  
A:Accession: S51901  
A:Molecule type: DNA  
A:Residues: 1-432 <VAV>  
A:Cross-references: EMBL:248149; NID:G663234; PIDN:CAA88164.1; PID:G663256  
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
Yeast 11, 1069-1075, 1995  
A:Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the  
a delta element.  
A:Reference number: S59156; MUID:96076631; PMID:7502582  
A:Accession: S59177  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-432 <VAV>  
A:Cross-references: EMBL:248149; NID:G663234; PIDN:CAA88164.1; PID:G663256  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995  
R:Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66791  
A:Accession: S66793  
A:Molecule type: DNA  
A:Residues: 1-432 <DUR>  
A:Cross-references: EMBL:274839; NID:G141947; PIDN:CAA99110.1; PID:G141948; GSPDB:GN00  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:WRS1; WRS1; MIPS:YOL097C  
A:Cross-references: SGD:S0005457  
A:Map position: 15L  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C:Keywords: ligase  
F:117-120/Region: ATP-binding motif (HXGH)

Query Match 55.0%; Score 1163; DB 2; Length 432;  
Best Local Similarity 54.8%; Pred. No. 5e-87;  
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEDFVDPWT-OTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQPHHFF 61  
19 STDVKEQVTPMDVGVDEQGAQNIIDYDKLIKQFTKPVNEETLKRFPQVGTGREPHF 78  
62 LRGGIFFSHRDKNQVLDAYENKPFYLTGRGSSSAMHVCHLIPFIYTKLQDVFNVL 121

Db 79 LRKGLFFSERDFTKILDVYEQGGKFPFLYTGSGPSSDSMHZGHMIPFVFTKWLQEVDFVPL 138  
QY 122 VIQMTDDEKYLWK-DLTLCAYGDAVENAKDIITACGFOINKTIFESDLDVYMGSSGYKN 180  
Db 139 VIETDDEKFLFKHUKITINDVKNFAENAKDIIVAGFDPKKTFFIMNSTYVG--GAFYET 196  
QY 181 VKIQKHVTFNQVKGIFGFTSDICIGKISPAIAQAASFNSPQIFRDRDIOCLIPCA 240  
Db 197 VVRVSQITGSTAKAVFGNDSICIGKEFASQIATAFPSSFPNVLGLPKDKTCLIPCA 256  
QY 241 IDDPYFRMTDVAIPRIGYPKAPALLHSTFFPALQCAQTKMSADPNSSIFLTDIAKQIKT 300  
Db 257 IDDPYFRMTDVAIPRIGYPKAPALLHSTFFPALQCAQTKMSADPNSSIFLTDIAKQIKT 300  
QY 301 KVNKHAFSGRDITIEHRQFGNCDVDSFMYLTFEEDDKLEQIRKDYTGSGAMLTGEL 360  
Db 317 KINKYAFSGGVADLHRELGGNPDVAVAYQFLSFFLDDDEKLAQLVNTYKAGTILSTGEM 376  
QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEEVT-PRKLS 395  
Db 377 KKLCIETLQEFVKAFQERRAQVDEETLCKFMVPHKL 412

RESULT 7  
C90190  
tryptophanyl-tRNA synthetase (trps) [imported] - *Sulfolobus solfataricus*  
C:Species: *Sulfolobus solfataricus*  
C>Date: 24-May-2001 #sequence revision 24-May-2001 #text\_change 15-Jun-2001  
C:Accession: C90190  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: *Sulfolobus solfataricus* complete genome.  
A:Reference number: A99139  
A:Accession: C90190  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-386 <KUR>  
A:Cross-references: GB:AE006641; NID:G13813608; PIDN:AAK40778.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: trps  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 42.9%; Score 907; DB 2; Length 386;  
Best Local Similarity 48.6%; Pred. No. 3.3e-66;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

QY 8 TEAEDFVDPWTQTSKAG-IDYDKLIVRFGSSKIDKELINRIERATGQPHHFFRGI 66  
Db 6 TMDPEFTVPWEV-----KGVVDYDKLIVQFGTKITEELKQRIKXLAGDL-HWMLRRV 59  
QY 67 FFSHRDKNQVLDAYENKPFYLTGRGSSAMHVCHLIPFIYTKLQDVFNVLQMT 126  
Db 60 FFSHRDKNQVLDAYENKPFYLTGRGSSAMHVCHLIPFIYTKLQDVFNVLQMT 119  
QY 127 DDEKYLWK-DLTLCAYGDAVENAKDIITACGFOINKTIFESDLDVYMGSSGYKNVYKIQ 185  
Db 119 DDEKYLWK-DLTLCAYGDAVENAKDIITACGFOINKTIFESDLDVYMGSSGYKNVYKIQ 175  
QY 186 KHVTFNQVKGIFGFTSDICIGKISPAIAQAASFNSPQIFRDRDIOCLIPCAIDQDP 245  
Db 176 KKLTFSEVRATFGDASSNIGLIFYPALQIAPT-----MFEKK---RCLIPAGIDQDP 225  
QY 246 YFRMTDVAIPRIGYPKAPALLHSTFFPALQCAQTKMSADPNSSIFLTDIAKQIKTKVNKH 305  
Db 226 YMLQDIDIAESLQYKAAQIHSKFLPLTGPESKSSSNPETAIVYDVPDPTVERKIMKY 285  
QY 306 AFSGGRTTIEHRQFGNCDVDSFMYLTFEEDDKLEQIRKDYTGSGAMLTGELKKAL 364  
Db 286 AFSGGGATIEHREKGNPDVAVAYQFLSFFLDDDEKLAQLVNTYKAGTILSTGEMGECI 345  
QY 365 IEVLQPLIAHQARRKEVTDEIVKEEVT-PRKLS 397



A: Experimental source: strain OT3  
A: Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C: Getatics:  
A: Gene: PH1921  
C: Superfamily: Yeast tyrosine-tRNA ligase  
C: Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 29.6%; Score 626.5; DB 2; Length 301;  
Best Local Similarity 44.2%; Pred. NO. 1.8e-43;  
Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

Cy 99 MGWGHLLPFTKVLQDVNPLVLTQMTDDEKYLWKD-LTLQAYGDAVENAKDIIACGF 157  
Db 1 MHGHIIPFPATKVLQKQEGNLYIQITDDEKFLPKENLTFTDTRKWAYDNILDIIVAGF 60  
Cy 158 DINTKTFSDUDYNGMSSGFYKNWVKIQHVTFNQVKGIFGTSCDICKISFPATQAP 217  
Db 61 DPDKTFIFQNSEF---TKIYEMAIPIAKKINFSMAKAVGFTQSKIGMIFPAIQIAP 116  
Cy 218 SFSNSFQIFRDRTDIOCLICADIDODPVFRTDVAPIGYPKPALHLESTFEPALOGAQ 277  
Db 117 TF-----PERK---RCILPAADIDDPYWRQLQDPAESLGIYKTAALKSKFVPSUTLS 166  
Cy 278 TMSGASDPNSSIFLDTAKQIKTKYNKHAFSGGRTIEHRQFGGNCDDVDVFMYLTFEL 337  
Db 167 GKMSASKPETALYLTDSFEDVEKKVWKETLTGGRTFLKQREKKGKGPVCKVFKWLEIFF 226  
Cy 338 EDDKLEQIRKDY---TSGMLTGBLKALILEVLQPLTAHQARRKEVTDIVKPEWTER 394  
Db 227 BEDDK--KLKERYACKNGBLTGCECKRYLSKIQEFLAEHQRRKK-AEKLVEKFKYTG 283  
Cy 395 KLS 397  
Db 284 KIA 286

RESULT 10  
384373  
tryptophanyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1  
C: Species: Halobacterium sp. NRC-1  
C: Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C: Accession: G84373  
E: Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Macdocks, D.G.; Jablor  
Gung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A: Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Liu  
A: Title: Genome sequence of Halobacterium species NRC-1.  
A: Reference numbers: A54160; MUID:20504483; PMID:11016950  
A: Accession: G84373  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-380 <STO>  
A: Cross-references: GB:AE004437; NID:glc0581646; PIDM:AAG20355.1; GSPDB:GN00138  
C: Getatics:  
A: Gene: trpS2  
C: Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 25.3%; Score 534.5; DB 2; Length 380;  
Best Local Similarity 35.4%; Pred. No. 8e-36;  
Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

Cy 10 ABEFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRPHFRLRGITFS 69  
Db 3 AGDNVDTPYAVESDD---LJYEKLAREGAEELTDQARFP-----DHP-VNRLGLFYA 53  
Cy 70 HRDMNQVLDAENKKEFYLYTGRGSSSEAMVGHILPIFTKWLQDVFNVPZVICTDDE 129  
Db 54 GRDVDDELTAEQS-----IVTGVPSG-PMLGHANVYFARLQDEFQATYVPLSDDE 108  
Cy 130 KYLWKDLTLQAYGDAVE-NAKDIACGFDINKTFIF---SLLXY-GMSSGFYKXVVKI 184  
Db 109 KYWFKDQTPAET-GQYLRANRDLIIAVGDFDELTRIVVDTRADVLYPLATFNGSV---164

185 QXHTVFNQVKGIFGFTDSCIGKISPPAIAQAPSFNSPFIQPRDRDIOCLIPCAIDQD 244  
165 -RHATLQNVG-----EPDNVGAFFPAVQAKL-L---LPQLVHG--EHETLPIAVDQD 213  
245 PYRMTTRDVAIRIGYP--KPALLHSTFFPALQOATKMSADPNSSIFLTDITAKQIKTKV 302  
214 PHRVSRDVAAKARYPVGKPCALLMQE-PSLAG-PGKMSSS-AGVSRLTDSPTVREKV 271  
303 NKHAFSGRDTIEHROFGNCVDVSEMYLTFFLEDD-KLEQIRKDYTSGLMTGELX 361  
272 RTHAYTGRASVSEHRKGVPAEDWQFQLSFAFFEDDAELAKIEREYRAGDILLSGELK 331  
362 KALIEVLQPLIAEHQARKEVTDIEVKEFTPRKLSFD 399  
332 DLAAADRITEFLAAHQRRALGD-VTEALDAFRLTDD 167

RESULT 11  
F64476  
tryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 03-Jun-2002  
C:Accession: F64476  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
R.; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.E.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: F64476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-370 <SUL>  
A:Cross-references: GB:U67582; GB:L77117; NID:gl592064; PIDN:AAB99425.1; PID:gl592065; T  
C:Genetics:  
A:Map position: FOR1375885-1376997  
A:Start codon: GTG  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 19.4%; Score 409.5; DB 2; Length 370;  
Best Local Similarity 30.9%; Pred. No. 1.2e-25;  
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
17 PWTVOTSSAKGIDYDKLIVRFSSSKIDKELINRIERATGQPHFLRRGIFPFSHRDMNQV 76  
8 PW-ETPAV--IDYKTMBOFGVKPIVWZGDLKEE-----HHFFRNILGHRDPERI 57  
77 LDAYENKKPFYLYTGRGSPSEAMHVGHLIPFIETKWLQ---DVFNVLVW-QMTDDEKYLW 133  
58 VDAIKNKKEFAVVGSMPPSK-MHFGHKWVVDLLKFQKTYDINIPi-----ADLEAYWA 112  
134 KDLTLQAYGDAV-ENAKDIIACGFBINKTFFSDLDYMGSSGFYKRWKVI-QKHVTEN 191  
113 RNMSFTTELALNEVITNYIALGLDPEKINLVYLSQYKQV-----KDLALILSKRTWS 167  
192 OVKGIFGFTDSCIGKISPPAIAQAPSFNSPFIQPRDR--DIQCLIPCAIDQDPPFM 249  
168 EMKAIYFGKSTNIGHVFAPIVQVADIL---HPQLDENLSPKPVVFPVGVIGDQPHRL 224  
250 TRDVAPR---IGYPKPLLHSTFFPALQOATKMSADPNSSIFLTDITAKQIKTKVAKHA 306  
225 TRDIANAKKFKTIPSSSTYHRWTGLGG--KMSSSKPTAFLTDDEKTVKKIIS-A 281  
307 FSGGRDTIEHROFGG---NCDVNVSEMYLTFFLEDDKLEQIRKDYTSGLMTGELXKA 363  
282 KTGGRTELEHKYGVGPEECWVVEFLY--HLILDDKELAEIYQKCRSGELTGCKCKM 339  
364 LIEVLQPLIAEHQARKEVTDIEV 388

340 AYERWVEFLKOLKEKREQAIEAVK 364

RESULT 12  
B69131  
tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Jun-2002  
C:Accession: B69131  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F  
; Qiu, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: B69131  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <MTH>  
A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84757.1; PID:g2621301  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH251  
A:Start codon: TTG  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.8%; Score 397.5; DB 2; Length 364;  
Best Local Similarity 27.9%; Pred. No. 1.1e-24;  
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;  
15 VDPWTVOTSSAKGIDYDKLIVRFSSSKIDKELINRIERATGQPHFLRRGIFPFSHRDMN 74  
2 IDPW----GSAK-LEYQDLLENFGVRPF-SEVLDEV-----PEFSWLMRGLIIFGHRDYE 50  
75 QVLDAVENKKPFYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVLVPIQMTDDEKYLW 134  
51 RIISAMKKGEDFAVVTGMFSGR-MHIGHKWIVDQLRW-YDRMGAEFIPIADMEAYSAR 108  
135 DLTLQAYGDAVEN-AKDIITACGFDNK-----TFIFSLDYMGSNGYKVVK 183  
109 GVDPEDSRR-ALIEYTAGYTAGLGLDEKDN-HVYLOSENLMLVEDLAYV----- 156  
184 IQKHVTFNWKVIGFTDSCIGKISPPAIAQAPSFNSPFIQPRDRDIOCLIPCAIDQ 243  
157 LAGKVNENELRALYGTGST-SMAHYAPIIQVSDILHPQLDELGGPR---PVIVVGPDQ 213  
244 DPYFMTROVAPRI---GYPKALLHSTFFPALQOATKMSADPNSSIFLTDITAKQIK 299  
214 DPHRLTRDIAARFDRYGFILPSSTYHREMGGLTGG--KMSNRPKSAIFLSDTPEAE 271  
300 TKYNKAFSGRDTIEHROFGNCVDVSEMYLTFFLE-DDKLEQIRKDYTSGLMTG 358  
272 AKI-RNAKTGRETLEKQRELGVPEECIIYETLLXHMSSGDSKLEELIYESCRNGILMCG 330  
359 ELKKALIEVLQPLIAEHQARKE 381  
331 ECKNTAEPIKPFEEELSVREK 353

RESULT 13  
T43806  
tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)  
C:Species: Encephalitozoon cuniculi  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 03-Jun-2002  
C:Accession: T43806  
R:Peyretailade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.  
Mol. Biol. Evol. 15, 693-699, 1998  
A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene  
A:Reference number: 222693; MUID:98277683; PMID:9615449  
A:Accession: T43806  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 &lt;PEY&gt;

A:Cross-references: EMBL:AJ012470; PIDN:CAA10034.1

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:keywords: ligase

Query Match 18.2%; Score 386; DB 2; Length 134;  
 Best Local Similarity 51.5%; Pred. No. 2.4e-24;  
 Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 12 EDFVDPVTVOTSSAK---GIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIF 67  
 Db 3 EORTPMDVEVVSDEVPVADYDKILNQCEKFNQALDRLEKLSGKPAHYFRRGIV 62  
 QY 68 FSHRDMNQVLDAYENKKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVPVLIQMTD 127  
 Db 63 FAHRDENLLDEIANNRRFYLTGRGSSKTHIGHTIPFLCLKYQDAFKRLVQITD 122  
 QY 128 DEKYLKDLTLD 139  
 Db 123 DEKFLWKSMLJE 134

RESULT 14

E:9461

t: tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C:Accession: B69461

R: Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

: Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

: Hloddek, A.; Zhou, L.; Overbeek, L.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, R.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98C49343; PMID:9389475

A:Accession: B69461

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 &lt;KLE&gt;

A:Cross-references: GB:AE000782; NID:92689309; PIDN:AAB89554.1; PID:9264885

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 17.5%; Score 370.5; DB 2; Length 420;  
 Best Local Similarity 27.3%; Pred. No. 2.2e-22;  
 Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 15 VDPMTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMN 74  
 Db 3 VTPMEVEGV-----IDYSKLIEFGMQPF-SBVLPEID-----NPHILMRGAIFGHRDYN 52  
 QY 75 QVLDAYENKKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVPVLIQMTDDEKYLWK 134  
 Db 53 RLIEAMQKPEPWAWSGFMPSG-LPHFGKMTMDLIVHMQSAGKAFV-ATADMEAHSVR 110  
 QY 135 DLTLUDQAYGDVAENAKIACGPDINKTFTFSD-LYMGSSGFGYKNVV-KIQKHVTFNQV 193  
 Db 111 GLSWEKTRGLMGLYKISIIALGLREDAVIYFQS-----KSHVKDLAFELSAENVFSEL 164  
 QY 134 KGIFGFTSDICIGKISFPAIQAPSFNSFPQIFDRDTDIQCLIPCAIDQDPYFMTDV 253  
 Db 165 RAIYGFNSDTSLAKMFWTAIQAADIL---HPQLSDFGFKPVPVPGADQDPHMLTRDL 221

QY 254 APRI----- 257  
 Db 222 AARLISFSEFPEVGVRVSRKGAELYSLSLDEEDKKIYEHMDFGEAREIERAVRKI 281  
 QY 258 -----GYPKALLHSTFFPALQGAOTKMSASDPSNSIFLTDTAKQIKTKVKNHAFSGG 310  
 Db 282 EVRIGGFAPIDPSSNYHRFTTGLTGG--KMSSKKEESVISLLDPPEGAKVMK-AFTGG 338  
 QY 311 RDTIEHRQFGNCVDVSYFMYLFFLED--DKLEIQIRKDYTSGLMILGELKXALIEVLQ 369

Db 339 RATAEQRRLGCEPDRCVVFELYSPHLIDSDDELNQIARCEGRLLCGCKKMAAELVK 398  
 QY 370 PLIAHQARRKEV 382  
 Db 399 SFLKEHQERMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1;

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jic-no, K.; Takaha

wa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10362966

A:Accession: D72477

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 &lt;KAM&gt;

A:Cross-references: DBJ:AP000064; NID:95105945; PIDN:BAA81476.1; PID:95106165

A:Experimental source: strain K1

C:Genetics:

C:Gene: APE2461

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.7%; Score 353; DB 2; Length 374;  
 Best Local Similarity 30.4%; Pred. No. 5e-2;  
 Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 15 VDPMTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMN 74  
 Db 8 LDPW---GAVEIKDYDKLRTFGIRPSEVL--PLLKAGWEPSELPRGIIFGHRDFD 61  
 QY 75 QVLDAYENKKPFYLTGRGSSSEAMHVGHLIPFTKWLQ-DVFNVELVIQMTDDEKYLWK 133  
 Db 62 KILEAKARGERVAVLTGFMPGSK-FHFGHKLTVQDLIYLQKNGFV--FVAIADAEAFV 118  
 QY 134 KDLTLUDQAYGDVAEN-ACQIACGPDINKT-FIPSDLYMGSSGFGYKNVVQKHVTFN 191  
 Db 119 RRIQEEAVRIAVEEYIANMIALGLDQDKTEYFQ---TNRGTPYFRLIQLFSGKVTAA 174  
 QY 192 QVKGIFG-FTSDICIGKISFPAIQAPSFNSFPQIFDRDTDIQCLIPCAIDQDPYFMT 250  
 Db 175 EMEALYGLTAPKXWASIT-----QAADILHVQLDVGGR---HVVVPVGAQDPHRLIT 227

QY 251 RDVAEP-----IGYKPAALLHSTFFPALQGAOTKMSASDPSNSIFLTDTAKQIKTKVKNHA 306  
 Db 228 RDLADRMAGVVELERFASTYHKLQPLDG--RKMSSSRSDSIFLTDPPEVARNKLFR-A 284  
 QY 307 FSGGRDTTIEHRQFGNCV-DVSPMYLTFLEDDDKLEIQIRKDYTS---GAMLTEGLKK 362  
 Db 285 ITGGRTAEEQRRLGGVPEVCVWEMDLYHMLPDDGVEVKHI---X--SRLGKILGCEQK 341  
 QY 363 ALIEVLQELIAHQARRKEVTDIEVKESTPR 394  
 Db 342 IAWEKLERFLAHHQSRLEKAKTIANKVNEPPR 373

Search completed: August 24, 2004, 18:03:46

Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:33:15 ; Search time 25 seconds

(without alignments)  
835.205 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNRGPDATEAEEDFVDPWTV.....VTDEIVKFMTPKLSGDFQ 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2101	99.3	471	1 SYW_HUMAN	P21381 homo sapien
2	2028.5	95.9	475	1 SYW_BOVIN	P17248 bos taurus
3	1961.5	92.7	475	1 SYW_RABIT	P23612 cryptolepus
4	1948	91.6	481	1 SYW_MOUSE	P12921 mus musculus
5	1210	57.2	395	1 SYW_SCHPO	Q09692 schizosacch
6	1163	55.0	432	1 SYW_YEAST	Q12109 saccharomyc
7	910.5	43.0	381	1 SYW_SULTO	Q976m1 sulfolobus
8	905	42.8	380	1 SYW_SULSO	Q972x0 sulfolobus
9	826.5	39.1	385	1 SYW_PYRPU	Q8u453 pyrococcus
10	803	37.9	385	1 SYW_PYRAB	Q9uy11 pyrococcus
11	796	37.6	386	1 SYW_PYRHO	Q59584 pyrococcus
12	735	34.7	375	1 SYW_PYRAE	Q8ztu5 pyrobaculum
13	534.5	25.3	380	1 SYW2_HALN1	Q9bn66 halobacteri
14	453	21.4	374	1 SYW_METKA	Q8tyf7 methanopyru
15	409.5	19.4	370	1 SYW_METTA	Q58810 methanococc
16	397.5	18.8	364	1 SYW_METTH	Q26352 methanobact
17	386	18.2	134	1 SYW_ENCCU	Q96771 encephalito
18	370.5	17.5	420	1 SYW_ARCPU	Q28579 archaeoglob
19	353	16.7	374	1 SYW_AERPE	Q9v924 aeropyrum p
20	329.5	15.6	437	1 SYW_METAC	Q8tual methanobac
21	313.5	14.8	491	1 SYW_METMA	Q9pwv5 methanobac
22	299.5	14.2	426	1 SYW_THEVO	Q978y8 thermoplasm
23	269.5	12.7	513	1 SYW1_HALN1	Q9hn83 halobacteri
24	268.5	12.7	426	1 SYW_THEAC	Q9hiw5 thermoplasm
25	203.5	9.6	341	1 SYW_STR3A	Q8e215 streptococc
26	203.5	9.6	341	1 SYW_STR3S	Q8dwp7 streptococc
27	192	9.1	323	1 SY1_ARCPU	Q29497 archaeoglob
28	185.5	8.8	341	1 SYW_STRPN	Q97n42 streptococc
29	183.5	8.7	340	1 SYW_STRP3	Q8ny22 streptococc
30	182.5	8.6	340	1 SYW_STRPY	Q99xh4 streptococc
31	181	8.6	340	1 SYW_STRMU	Q8dr11 streptococc
32	175.5	8.3	341	1 SYW_CLOJO	Q46127 clostridium
33	174.5	8.2	341	1 SYW_LACLA	Q9c5d1 lactococcus

34 172 8.1 394 1 SYW\_YEAST  
35 170.5 8.1 366 1 SYW\_SULSO  
36 169 8.0 395 1 SYW\_AQUAE  
37 166.5 7.9 346 1 SYW\_GILTR  
38 163.5 7.7 353 1 SYW\_BORBU  
39 159.5 7.5 337 1 SYW\_TREPA  
40 158 7.5 346 1 SYW\_CHLMC  
41 154 7.3 345 1 SYW\_CHLCV  
42 152 7.2 344 1 SYW\_CHLPN  
43 146 6.9 326 1 SYW\_HELFP  
44 142.5 6.7 337 1 SYW1\_STR2M  
45 138 6.5 319 1 SYW\_METTH

## ALIGNMENTS

RESULT 1  
SYW\_HUMAN  
ID SYW\_HUMAN STANDARD; PRT; 471 AA.  
AC P23381; P78535; Q9UDL3;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptostan--tRNA ligase)  
DE (TIPRS) (IFP53) (HWR5).  
GN WARS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OK NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92105071; PubMed=1761529;  
RA Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;  
RT "Interferon induces tryptophanyl-tRNA synthetase expression in human  
RT fibroblasts.";  
RL J. Biol. Chem. 266:24245-24248(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92107982; PubMed=1763065;  
RA Fleckner J., Rasmussen H.H., Justesen J.;  
RT "human interferon gamma potentially induces the synthesis of a 55-kDa  
RT protein [gamma 2] highly homologous to rabbit peptide chain release  
RT factor and bovine tryptophanyl-tRNA synthetase.";  
RN Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92112058; PubMed=1765274;  
RA Frolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinzivieva O.L.,  
RA Kisselev L.L.;  
RT "Cloning and nucleotide sequence of the structural gene encoding for  
RT human tryptophanyl-tRNA synthetase.";  
RL Gene 109:291-296(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92164636; PubMed=1537332;  
RA Buwitt U., Flohr T., Boettger E.C.;  
RT "Molecular cloning and characterization of an interferon induced  
RT human cDNA with sequence homology to a mammalian peptide chain  
RT release factor.";  
RN EMBO J. 11:489-496(1992).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins J., Wagner J., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Xooe T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.Y., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carrinci P., Frange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., McIlahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." [6]  
Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903 (2002).

SEQUENCE OF 1-13 FROM N.A.  
MEDLINE=9631994; PubMed=8724762;  
Sokolova I.V., Narovlianskii A.N., Achenkova A.M., Turpaev K.T.; "Alternative splicing of 5'-terminal exons of the human tryptophanyl-tRNA synthetase gene." [7]  
Mol. Biol. (Mosk) 30:319-329 (1996).

SEQUENCE OF 1-141 AND 182-471 FROM N.A.  
TISSUE-Sperm;  
MEDLINE=93292992; PubMed=7685728;  
Frolova L.Y., Grigorieva A.Y., Sudomoina M.A., Kisselev L.L.; "The human gene encoding tryptophanyl-tRNA synthetase: interferon-response elements and exon-intron organization." [8]  
Gene 128:237-245 (1993).

SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365;  
TISSUE-Keratinocytes;  
MEDLINE=93162043; PubMed=1286667;  
Rasmussen H.H., van Damme J., Fuype M., Gesser B., Celis J.E., Vandeckerckhove J.; "Microsequences of 145 proteins recorded in the two-dimensional gel protein database of normal human epidermal keratinocytes." [9]  
Electrophoresis 13:960-969 (1992).

FUNCTION.  
MEDLINE=92225128; PubMed=1373391;  
Bange P.-C., Flohr T., Buwitt U., Boettger E.C.; "An interferon-induced protein with release factor activity is a tryptophanyl-tRNA synthetase." [10]  
FEBS Lett. 300:162-166 (1992).  
-!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP + diphosphate + L-tryptophanyl-tRNA(Trp).  
-!- SUBUNIT: Homodimer.  
-!- INDUCTION: By interferon gamma.  
-!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
-!- SIMILARITY: Contains 1 WHEP-TRS domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M77804; AAA67324.1; -  
EMBL; X59892; CAA42545.1; -  
EMBL; M61715; AAA61298.1; -  
EMBL; X62570; CAA44450.1; -  
EMBL; BC017489; AAL17489.1; -  
EMBL; S82905; AAB39381.1; -  
EMBL; X67920; CAB94198.1; -  
EMBL; X67921; CAB94198.1; JOINED.  
EMBL; X67922; CAB94198.1; JOINED.  
EMBL; X67923; CAB94199.1; -  
EMBL; X67924; CAB94199.1; JOINED.  
EMBL; X67925; CAB94199.1; JOINED.  
EMBL; X67926; CAB94199.1; JOINED.  
EMBL; X67927; CAB94199.1; JOINED.

DR EMBL; X67928; CAB94199.1; JOINED.  
DR PIR; A41633; A41706.  
DR Aarhus/Ghent-2DPAGE; 3524; IEF.  
DR PHCI-2DPAGE; P23381; -.  
DR Genew; HGNC:12729; WARS.  
DR MIM; 191050; -.  
DR GO; GO:0005737; C-cytoplasm; TAS.  
DR GO; GO:0005625; C-soluble fraction; TAS.  
DR GO; GO:0002885; P-negative regulation of cell proliferation; TAS.  
DR GO; GO:0008412; P-protein biosynthesis; TAS.  
DR GO; GO:0006436; P-tryptophanyl-tRNA aminoacylation; TAS.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR002306; tRNA-synt\_1b.  
DR InterPro; IPR002306; tRNA-synt\_1b.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR Pfam; PF00458; WHEP-TRS; 1.  
DR PRINTS; PRO1039; TRNASYNTHTRP.  
DR TIGRFAMs; TIGR00233; tips; 1.  
DR PROSITE; PS02178; AA TRNA LIGASE\_1; 1.  
DR PROSITE; PS02762; WHEP-TRS; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
FT DOMAIN 19 64 WHEP-TRS.  
FT SITE 164 173 "HIGH" REGION.  
FT SITE 349 353 "KMSKS" REGION.  
FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
FT CONFLICT 424 424 A -> R (IN REF. 4).  
SQ SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;

Query Match 99.3%; Score 2101; DB 1; Length 471;  
Best Local Similarity 99.5%; Pred. No. 2.3e-163;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGPDATAEAEEDFVDPWTQVTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60  
DB 71 SNHGPDATAEAEEDFVDPWTQVTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 130  
QY 61 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRPSSEAMHVGHLIPFTFKWLDQVENV 120  
DB 131 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRPSSEAMHVGHLIPFTFKWLDQVENV 190  
QY 121 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIACGFDINKTFFISDLDYMGSSSFYKN 180  
DB 191 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIACGFDINKTFFISDLDYMGSSSFYKN 250  
QY 181 VVKIQKHVTENQKVGIFGFTSDCGIKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 240  
DB 251 VVKIQKHVTENQKVGIFGFTSDCGIKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 310  
QY 241 IDQDPYFRMTDVAIPRGYKPKALLHSTFFPALQAGTQKMSASDNNSSIFLDTAKQIKT 300  
DB 311 IDQDPYFRMTDVAIPRGYKPKALLHSTFFPALQAGTQKMSASDNNSSIFLDTAKQIKT 370  
QY 301 KVNKHFSGGRTTIEHRQFGNCDDVSVFMYLTFLEDDKLEQIRKQYTSGLMTGEL 360  
DB 371 KVNKHFSGGRTTIEHRQFGNCDDVSVFMYLTFLEDDKLEQIRKQYTSGLMTGEL 430  
QY 361 KKALIEVLQPLIAEHQARKKXVTHIVKEFMTPRKLSDFQ 401  
DB 431 KKALIEVLQPLIAEHQARKKXVTHIVKEFMTPRKLSDFQ 471

## RESULT 2

SWW\_BOVIN  
ID SWW\_BOVIN STANDARD; ERT; 475 AA.  
AC P17248;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpRS).  
GN WARS.  
OS Bos taurus (Bovine).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina;  
 RX MEDLINE=91329348; PubMed=1907847;  
 RA Garrett M., Pajot B., Trezequet V., Labouesse J., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sallatranque M.-L., Alteriot J.,  
 RA Labouesse J., Bonnet J.;  
 RA Guenou M., Sarger C., Labouesse B., Bonnet J.;  
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 RT prokaryotic synthetases but near identity with mammalian peptide  
 RT chain release factor.";  
 RL Biochemistry 30:7809-7817(1991).  
 RN [2]  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Garrett M., Trezequet V., Pajot B., Gandar J.-C., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sarger C., Alteriot J., la Boussec B.,  
 RA Labouesse J., Bonnet J.;  
 RA Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 RL -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AYP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X53918; CAA37872.1; -  
 DR EMBL; X52113; CAA36356.1; -  
 DR F1R; A40279; YWBO.  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1.  
 DR InterPro; IPR002306; Trp tRNA-synt 1b.  
 DR InterPro; IPR000738; WHEP-TRS.  
 DR Pfam; PF00579; tRNA-synt 1b; 1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS02178; AA TRNA\_LIGASE\_I; 1.  
 DR PROSITE; PS02762; WHEP-TRS; 1.  
 K1 Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 F1 DOMAIN 24 69  
 F2 DISPENSABLE TO THE CATALYTIC ACTIVITY.  
 F3 SITE 1 117  
 F4 SITE 169 178  
 F5 SITE 353 357  
 F6 "KMSKS" REGION.  
 F7 CONFLICT 17 17 L -> M (IN REF. 2).  
 S2 SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 95.9%; Score 2028.5; DB 1; Length 475;  
 Best Local Similarity 95.8%; Pred. No. 1.9e-157;  
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;  
 Q/ 1 SNHGPDAAEEDFVDPMTVQTSSAGKIDYDKLVRFGSSKIDKELINRIERATQORPH 60  
 D/ 76 SGEGLDAEADDFVDPMTVQTSSAGKIDYDKLVRFGSSKIDKELINRIERATQORPH 135  
 Q/ 61 FLRGIFPSHRMNVQLDAYENKPFYLYTGSPSSSAHVGHLLPFYFTKWLQDVENV 120  
 D/ 136 FLRGIFPSHRMNVQLDAYENKPFYLYTGSPSSSAHVGHLLPFYFTKWLQDVENV 195  
 Q/ 121 LVIQMTDDEKYLWDLTDQAYGDAVENAKOTIACGFDINKTFIFSDLDYMGMSGFYKN 180  
 D/ 196 LVIQMTDDEKYLWDLTDQAYGDAVENAKO-ITCGFDINKTFIFSDLDYMGMSGFYKN 254

QY 181 VVKIQHVTNFQVKIGFPGTSDSCIGKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 240  
 DB 255 VVKIQHVTNFQVKIGFPGTSDSCIGKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 314  
 QY 241 IDODPYERWTRDVAIRIGYKPKALAHSTFFPALQGAOTKWSASDPSNIFLDTCAKQIKT 300  
 DB 315 IDODPYERWTRDVAIRIGYKPKALAHSTFFPALQGAOTKWSASDPSNIFLDTCAKQIKT 374  
 QY 301 KYNKHAFIGSGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 360  
 DB 375 KYNKHAFIGSGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 434  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSYDFQ 401  
 DB 435 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSYDFQ 475  
 RESULT 3  
 SWW\_RABIT  
 ID SYW\_RABIT STANDARD; PRT; 475 AA.  
 AC P23612; Q28607;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)  
 DE (TPRS).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90239043; PubMed=2185472;  
 RA Lee C.C., Craig W.J., Muzny D.M., Harlow E., Caskey C.T.;  
 RT "Cloning and expression of a mammalian peptide chain release factor  
 RT with sequence similarity to tryptophanyl-tRNA synthetases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).  
 RN [2]  
 RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.  
 RX MEDLINE=94009008; PubMed=8404867;  
 RA Prolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,  
 RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-J.;  
 RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA  
 RT synthetase are distinct proteins.";  
 RL EMBL J. 12:4013-4019(1993).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBUNIT: Homodimer (by similarity).  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.  
 CC -!- CAUTION: Was originally (Ref.1) thought to be a eukaryotic release  
 CC factor (BRP).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M33460; AAA31246.1; ALT SEQ.  
 DR EMBL; U02595; AAB60257.1; -  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1.  
 DR InterPro; IPR002306; Trp tRNA-synt 1b.  
 DR InterPro; IPR000738; WHEP-TRS.  
 DR Pfam; PF00579; tRNA-synt 1b; 1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS00178; AA TRNA\_LIGASE\_I; 1.



```
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;
Query Match 92.7%; Score 1961.5; DB 1; Length 475;
Best Local Similarity 91.5%; Pred. No. 5.3e-152;
Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;
QY 2 NHG-PDATAEEDFDVDPWTVTQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60
DB 75 SHGDPEAVDDKEDFDVDPWTVRTSSAKGIDYDKLIVRFGSSKIDKELVNRIERATGQRPHR 134
QY 61 FRRGGTFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMHVGHILIPFIETKWLQDVFNVP 120
DB 135 FRRGGTFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMHVGHILIPFIETKWLQDVFNVP 194
QY 121 LVQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 180
DB 195 LVVQMSDDEKYLWKDLTLDOAYGYTLLENAKDIIACGFDVKNKTFIFSDLDYMGSPGFYKN 254
QY 181 VVKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCA 240
DB 255 VVKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCA 314
QY 241 IQDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKT 300
DB 315 IQDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKT 374
QY 301 KYNKHAFGSGRDTIEHRFGGNCNDVDSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 360
DB 375 KYNKHAFGSGRDTIEHRFGGNCNDVDSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 434
QY 361 KALIEVLOPLIAEHQARRKEVTDIVKEFMTPKLSFDFQ 401
DB 435 KXELIDVLOPEVAEHQARRKEVTDIVKEFMTPKLSFDFQ 475
DI
```

## RESULT 4

```
SYW_MOUSE
III SYW_MOUSE STANDARD; PRT; 481 AA.
AC P12921;
DB 01-OCT-1993 (Rel. 27, Created)
D7 01-OCT-1993 (Rel. 27, Last sequence update)
D7 28-FEB-2003 (Rel. 41, Last annotation update)
DI Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DI (TPRS).
DI WARS OR WRS.
DI Mus musculus (Mouse).
DI Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DI Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DI NCBI_TaxID=10090;
DI [1]
DI SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
DI MEDLINE=95018226; PubMed=7932716;
DI Pajot B., Sarger C., Bonnet J., Garret M.;
DI "An alternative splicing modifies the C-terminal end of tryptophanyl-
DI tRNA synthetase in murine embryonic stem cells.";
DI J. Mol. Biol. 242:599-603(1994).
DI -! CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
DI diphosphate + L-tryptophanyl-tRNA(Trp).
DI -! SUBUNIT: Homodimer (By similarity).
DI -! ALTERNATIVE PRODUCTS.
DI Event=Alternative splicing; Named isoforms=2;
DI Name=1; Synonyms=long;
DI IsoId=P12921-1; Sequence=Displayed;
DI Name=2; Synonyms=short;
DI IsoId=P12921-2; Sequence=VSP_006313;
DI -! TISSUE SPECIFICITY: Isoform 2 is widely expressed, isoform 1 is
DI found only in embryonic stem cells.
DI -! SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
```

```
CC -! SIMILARITY: Contains 1 WHEP-TRS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69656; CAA49347.1; -.
DR EMBL; X69657; CAA49348.1; -.
DR PIR; S50053; S50053.
DR MGD; MGI:104630; MARS.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001432; tRNA-synt_1.
DR InterPro; IPR002306; Trp_tRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRfam; TIGR00233; tRps; 1.
DR PROSITE; PS00178; AA-tRNA_LIGASE_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPLIC 476 481 Missing (in isoform 2).
FT FTID=VSP_006313.
SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;
Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 4.4e-150;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;
QY 2 NHGPDATAEEDFDVDPWTVTQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61
DB 76 NCSDATKASEDFVDPWTVTQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 135
QY 62 LRRGIFPSHRDMNQVLDAVENKPPVLYTGRGSSSEAMHVGHILIPFIETKWLQDVFNVP 121
DB 136 LRRGIFPSHRDMNQVLDAVENKPPVLYTGRGSSSEAMHVGHILIPFIETKWLQDVFNVP 195
QY 122 VIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 181
DB 196 VIQSDDEKYLWKDLTLDOAYGYTYVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 255
QY 182 VKIQEVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCAI 241
DB 256 VKIQEVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCAI 315
QY 242 DQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKTK 301
DB 316 DQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKTK 375
QY 302 VNKHAFSGRDTIEHRFGGNCNDVDSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 361
DB 376 VNKHAFSGRDTIEHRFGGNCNDVDSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 435
QY 362 KALIEVLOPLIAEHQARRKEVTDIVKEFMTPKLSFDFQ 401
DB 436 KTLIDVLOPLIAEHQARRKAVTEETVKEFMTPKLSFDFQ 475
DI
```

## RESULT 5

```
SYW_SCHPO
ID SYW_SCHPO STANDARD; PRT; 395 AA.
AC C09592;
DI 01-NOV-1995 (Rel. 32, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
```

Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpRS).

DE Schizosaccharomyces pombe (Fission yeast).  
GN SPAC2F7.13C.  
OS Schizosaccharomycetes; Schizosaccharomycetaceae;  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetaceae;  
CC Schizosaccharomycetes; Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]

SEQUENCE FROM N.A.

RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Williams K., Rajendram M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgourou J., Peat M., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy J., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkmer G., Aert R., Robben J., Grymonprez B.,  
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu S., Dreano S., Gloux S., Lelaure V., Mottier S.,  
FA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
FA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
FA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
FA Dominguez A., Revuelta J., Moreno S., Armstrong J., Farsburg S.L.,  
FA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
FA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
FJ "The genome sequence of Schizosaccharomyces pombe";  
FJ Nature 415:871-880(2002).

CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z50142; CAA90500.1; --  
DR PIR; S58157; S58157.  
DR GeneDB\_SPombe; SPAC2F7.13C; --  
DR InterPro; IPR002305; trna-synt\_1b.  
DR InterPro; IPR001412; trna-synt\_1.  
DR InterPro; IPR023306; Trp tRNA-synt\_1b.  
DR Pfam; PF00579; trna-synt\_1b; 1.  
DR PRINTS; PR01039; TRNASYNTHRP.  
DR TIGRFAMs; TIGR00233; trps; 1.  
DR PROSITE; PS00178; AA tRNA LIGASE I; 1.  
DR KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;  
FT Ligase; ATP-binding.  
FT SITE 91 100 "HIGH" REGION.  
FT SITE 275 279 "RMSKS" REGION.  
SQ SEQUENCE 395 AA; 44910 MW; E656AEBB76C5FDF9 CRC64;

Query Match 57.2%; Score 1210; DB 1; Length 395;  
Best Local Similarity 59.6%; Pred. No. 5.7e-91;  
Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

Qy 11 BEDFVDPWTQTS-----SAKGDYDKLIVRFSSKIDKELINIERATQRPHHFLRRGI 66  
Db 4 BEQIVTPDWKSGIVDGEKGDYELRVQFGRTKTPQLERFEKLTGKPKHLLRRGA 63

Qy 67 FESHEOMNOVLDAENKKPFYLTGROPSSSEAMVGHLLPEIFTKWLODVENVELVIMT 126  
Db 64 FESHRDFMLIDRYEQKKPFYLTGROPSSSDSHLGHMIPFMCKWLODVFEQVELVIMT 123  
Qy 127 DDEKYLWKD-LTLDQAYGDAVENAKDIIACGFDINKTFIESDLDYMGSSGFYKNVYKIQ 185  
Db 124 DDEKFLKQGVSLDCCORFARENAXDIIIVDFEKKTFIFWNSTYVG--GAFYQNVVRIA 181  
Qy 186 KHTVFQVKGIFGPTSDSGCKISFPALQAPZFSNSFPQIFRDRTPICQLPICAIDDP 245  
Db 182 KCITANGKACFGFTSDSGIKTHFASIQAAZPFSFSSFFHFNCAKIDPCLPCAIDQDP 241  
Qy 246 YFMTVDVABRIGYPKPALHSTFEPALOGACQKMSASDPSNSTFELDTAKCIKTKNKH 305  
Db 242 YFELTRDVSRLPKPKPALHSHRFPFALQGPQSKMSASKSSAIFMTDTPNKIKKINRH 301  
Qy 306 AFSGGRTDTEHHRQFGNGCDVDSFVMTLTFLEDDKLEQIRKDYTGAMLTGELKKALI 365  
Db 302 AFSGGGATTEHREKGNPDVDVAYQVLSFFLDDDEKLKLYNTYKAGTSLTGEMKGECI 361  
Qy 366 EVLQPLAEHQARKEVTEIVKEFMT-PRKLSF 398  
Db 362 KLLQQFVSDFQAARSKYDEATLDMFMDGSRKLEW 395

# RESULT 6

SYNC\_YEAST STANDARD; PRE; 432 AA.  
ID SYNC\_YEAST AC Q12183;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--  
DE tRNA ligase) (TrpRS).  
GN WRS1 OR YOL097C OR HRE432.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96076631; PubMed=7502582;  
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;  
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV  
RT including the Trp-H3 retrotransposon, the sufi(+) frameshift  
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a  
RT delta element";  
RL Yeast 11:1069-1075(1995).  
RN [2]  
RP FUNCTION  
RX MEDLINE=97197969; PubMed=9046085;  
RA John T.R., Ghosh M., Johnson J.D.;  
RT "Identification and expression of the Saccharomyces cerevisiae  
RT cytoplasmic tryptophanyl-tRNA synthetase gene";  
RL Yeast 13:37-41(1997)  
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z48149; CAA98164.1; --  
DR EMBL; Z74839; CAA99110.1; --  
DR PIR; S51901; S51901.

"Complete genome sequence of an aerobic thermoacidophilic

```

OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RA [1]
R2 SEQUENCE FROM N.A.
R3 STRAIN=ATCC 35092 / DSM 1617 / P2;
R4 MEDLINE=21332296; PubMed=11427726;
R5 She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
R6 Aweya M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
R7 De Moors A., Erasuo G., Fletcher A.C., Gordon P.M.K.,
R8 Haikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
R9 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
R10 Charlebois R.L., Doolittle W.F., Duguet X., Gaasterland T.,
R11 Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
R12 "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
R13 Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
R14
R15 -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
R16 diphosphate + L-tryptophanyl-tRNA(Trp).
R17 -I- SUBCELLULAR LOCATION: Cytoplasmic.
R18
R19 -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
R20
C1 This SWISS-PROT entry is copyright. It is produced through a collaboration
C2 between the Swiss Institute of Bioinformatics and the EMBL outstation-
C3 the European Bioinformatics Institute. There are no restrictions on its
C4 use by non-profit institutions as long as its content is in no way
C5 modified and this statement is not removed. Usage by and for commercial
C6 entities requires a license agreement (See http://www.isb-sib.ch/announcement/
C7 or send an email to license@isb-sib.ch).
C8
C9 EMBL; AE006677; AAK40778.1; ALT_INIT.
C10
C11 DR HAMAP; MF_00140; -; 1.
C12
C13 DR InterPro; IPR002305; tRNA-synt_1b.
C14
C15 DR InterPro; IPR001412; tRNA-synt_1.
C16
C17 DR InterPro; IPR002306; Trp tRNA-synt_1b.
C18
C19 DR Pfam; PF00579; tRNA-synt_1b; 1.
C20
C21 DR PRINTS; PR01039; TRNASYNTHTRP.
C22
C23 DR TIGRFAMS; TIGR00233; trps; 1.
C24
C25 DR PROSITE; PS00178; AA tRNA LIGASE I; FALSE NEG.
C26
C27 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
C28 Complete proteome.
C29
C30 SITE 81 89 "HIGH" REGION.
C31 FT SITE 253 257 "KMSKS" REGION.
C32
C33 SQ SEQUENCE 380 AA; 44691 MW; CF8344CF63083680 CEC64;
C34
Query Match 42.8%; Score 905; DB 1; Length 380;
Best Local Similarity 48.7%; Pred.No. 3.5e-66;
Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;
C35
QY 11 EEDFVDPWTVOTSSAKG-IDYDKLIVFGSSKTDKELINKRIERATGQRPHEFLESGIFFS 69
DB 3 DEFVTPEV-----KGVDYDKLIVGFGTKTELEKQRIKNLAGDL-HVLEKRNFFS 56
C36
QY 70 HRDMNQVLDAYENKKPFYLYTGRGPSSEAMHVGLIPIFTKMLQDVFNVELVIQMTDDE 129
DB 57 HRDLDLVLNDYKSGKGFPLYTGRAPSL-GMHIGHLIPIFTKMLQEKFNAMLYEITDDE 115
C37
QY 130 KYLWK-DLTLQAYGDAVENAKILACGPDINKTIFPSDLIDYMGMSGGFYKNVYKIQHV 188
DB 116 KYMRNPEFTLDQTRWAYDNILDIIVAGFNECKTFFIQDEYI---RNMYPIITVKIAKLL 172
C38
QY 189 TFMQVKGIFGFTSDCIKCK-SFPAIQAPSFNSFPQIFRDRTDIQCLPCALDQDEYFR 248
DB 173 TFSVVRATFGLDASSNGLIFYPALQAPT-----MFEKK---RCLIPAGIDQDDFYR 222
C39
QY 249 MTRDVPRIQYKPKALLHSTFFPALQGAQTKMGASDPNSSIFLTDTAQKTIKVNKHAFS 308
DB 223 LQRDIAESLGYYKAAQIHSKFLPLTGPCKMGSSNPETALYLVDDPKTVERKIMKYAFS 282
C40
QY 309 GGRDTTEHRQFGNCVDVDSFMYLIFLEDDD-KLEQIKNDYTSGLMFGELKALIEV 367
DB 283 GGOFTTELHKYGNPEIDPFQWLYYFFEDDNRIKEIEEYRSGKMLFGELKQILIDK 342

```

```

QY 368 LQPLIAEHQARKEVTDIEIVKGFMTPTKLS 397
      | : ||: ||: | : ||: | : ||: | : ||:
Db 343 LNNFLEHR-RRSEAEKELVHVFKYDGKLA 371

RESULT 9
SYW PYRFU STANDARD; PRT; 385 AA.
AC Q8U453;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
DE DE
GN TRPS OR PF0241.
OS Pyrococcus furiosus.
OS Archaea; Euryarchaeota; Thermococci; Thermococcaceae; Thermococcaceae;
OC Pyrococcus.
OC NCBI_TaxID=2261;
ON RN
ON RN
ON RN
SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RC Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RA "the complete sequence of the Pyrococcus furiosus genome.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ Databases.
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC -1- diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE010149; AAL80365.1; -
CC HAMAP; MF 00140; -; 1.
CC InterPro; IPR002305; tRNA-synt_1b.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002306; Trp_tRNA-synt_1b.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PR01039; TRNASYNTHTRP.
CC TIGRFAMS; TIGR00233; tips; 1.
CC PROSITE; PS00178; AA tRNA Ligase 1; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SITE 82 90 "HIGH" REGION.
CC FT SITE 253 257 "KMSKS" REGION.
CC SEQUENCE 385 AA; 45178 MW; 3A7A628958200CCC CRC64;

Query Match 39.1%; Score 826.5; DB 1; Length 385;
Best Local Similarity 45.9%; Pred. No. 8.e-60;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

QY 11 EEDF-YDPFVTQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATQRFHFLRGIFFS 69
Db 2 EBEFKVTEVEGI---IDYKKLIHQFGTSPITDLDLLEARTLTKSELPIFFRRKFFS 57
QY 70 HRDMNQVLDAYENKPFYLYTGRGSSSEAMHGHIPFFFTWLDQVFNVLVIQMTDDE 129
Db 58 HRDYDKVLDYDQGGFFLYTGRGSG-PMHGHIPFFATKWLQKQFVNLYIQITDDE 116
QY 130 KYLWKD-LTLDQAYGDAVENAKDITACGDFDINKTFTFSOLDYMGSSSGYKNVWIKQHV 188
Db 117 KFLFENLITFDTKWAYQNLIDITANGDPDKTIFQNSP-----TKIYEWAIPTAKKI 172
QY 189 TRNQVKGIFGFTSDSCIGKISFPAICAAPSFNSPPQIFEDRTDIQCLIPCAIDQDPYFR 248
Db 173 NFSMAKAVGFTFQSGKIGMIFPAICAAPTF-----FEKK---RLCPAIDQDPYWG 222


```

QY	71	RDMMQVLDAYENKAPFYLYTGRGSSSEAMVGHILPIFTKWLQDVKNVPLVIMQTDDEK	131
Db	59	RDVMDQLQDYEEGRGFFLYTGRGSG-PMHIGHIIPFPFATKWLQSKFGVNLVIQITDDEK	117
QY	131	YLWKK-LTLDAQVGDVAENAKDILACGFDINKTPIESDLDYMGMSGFEKNVVK-QKHVT	189
Db	118	FLFKENLTFEDTKHWAVENTLIDILAVGDPDKTIFQNSEP-TKYEMAIPTAKKIN	173
QY	190	FNQVKGIFGTFDSDICIGKISFPATQAAAPFSMSPPQIFRDRDTDIQCLIPCAIDQDPYFRM	249
Db	174	FSMAKAVGFGTEQSKIGMIFFPAIQIAPTF-----FEKR--RCLIPAAIDQDPYWL	223
QY	250	TRDVAPRIGYPKALLHSTFPALQAGQTKMSASDPNSSILFTDTAKOIKTKVNKHAFSG	309
Db	224	QRDFASLSGYYKTAHHSKVPVSLTSGKMSASKPETA-YLTDSPEDEVKVKWFALTG	283
QY	310	GRDTIEHRFGGNCVDVSPMYLTFLEDDDKLEQIRKDY--TSGAMLTGELKAKALIE	366
Db	284	GRPTLKEQREKGEPEKCVVFKWLEIFEEDDK--KLKERYACKNGELTGECKRYLIS	341
QY	367	VLQPLIAHQARKEVYDITVK 388	
Db	342	KICEFJKEHQRRKRAEKQIEK 363	

RESULT 11

SYN_PTHO	STANDAED;	PRT;	386 AA.
AC	059584;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Tryptophanyl-tRNA synthetase [EC 6.1.1.2] (Tryptophan--tRNA ligase)		
DE	(TrpAs).		
DE	TRPS OR PHI921.		
OS	Pyrococcus horikoshii.		
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;		
OC	Pyrococcus.		
OX	NCBI_TaxID=53953;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=OT3.		
RC	MEDLINE=9834137; PubMed=9679194;		
RA	Kawarabayashi Y., Sawada M., Horikawa H., Kaikawa Y., Hino Y.,		
RA	Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,		
RA	Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,		
RA	Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,		
RA	Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,		
RA	Masuchi Y., Shizuya H., Kikuchi H.;		
RA	"Complete sequence and gene organization of the genome of a hyper-		
RT	thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";		
RL	DNA Res. 5:55-76(1998).		
CCC	-!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +		
CCC	diphosphate + L-tryptophanyl-tRNA(Trp).		
CCC	-!- SUBCELLULAR LOCATION: Cytoplasmic		
CCC	-!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.		

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; AP000007; BAA31046.1; ALT_INIT.
CC	HMAP; MF 00140; -; 1.
DR	InterPro; IPR002105; tRNA-synt 1b.
DR	InterPro; IPR001412; tRNA-synt_1.
DR	InterPro; IPR002306; Trp tRNA-synt 1b.
DR	Pfam; PF00579; tRNA-synt 1b; 1.
DR	PRINTS; PR01039; TRNASYNTHTRP.





RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,  
 RA Alan M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,  
 RA "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE005109; AAG20355.1; --  
 CC PIR; G84373; G84373.  
 DR HAMAP; MF\_00140; --; 1.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00579; tRNA-synt\_1b.  
 DR PRINTS; PR01039; tRNA-synt\_1b.  
 DR TIGRams; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA-TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 74 82 "HIGH" REGION.  
 FT SITE 249 253 "KMSKS" REGION.  
 FT SITE 380 AA; 41936 MW; 0F7B2B95386404F CRC64;  
 SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B95386404F CRC64;  
 Query Match 25.3%; Score 534.5; DB 1; Length 380;  
 Best Local Similarity 35.4%; Pred. No. 4.9e-36;  
 Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;  
 QY 10 AEDFDVPTWTQSSAKGIYDKLIVRGSSKIDKELINRIERATGQPHHFLRGIFFS 69  
 DB 3 AGDNVTPVAVESDD---LDYKLLARFGADELTDQARFP-----DHPLVNRGLFYA 53  
 QY 70 HRDMQVLDAYENKPPFLYLTGRGPSSSEMHVGLIPEFTKWLQDVFNVLVIQMTDDE 129  
 DB 54 GRNVDLFLTAGES-----IVTGVGPGG-FMHLGHAMVFFPARLQDEFGARVYVPLSDDE 108  
 QY 130 KYLWKDLTLQDQAVGDAVE-NAKDIACGDFINKFIF---SLLDYM-GMSGSEYKXVVKI 184  
 DB 109 KYWFKDQTPAET-GDYLRLMLRDLAVGDPPELTIRIVDFRDADVLVPLATAGDV--- 164  
 QY 185 QKHVTFNQYKGFIFGFTSDICIGKISFPQAQPSFNSFPQIFRDRDIOCLIPCAIDQD 244  
 DB 165 -RHATLQNYG-----EPDNVGOAFYPAVQTALL---LPQLVHG--EHETIVPIAVDQD 213  
 QY 245 PYFRMTRDVAPRIGYP--KPALLHSTFFPALQCAQKMSASDPSNSIFLTDKAKIKTKV 302  
 DB 214 PHRVSRSDVAARYPVGPFGALLMQLFSLAG-PGKMSSS-AGVSIRLTDSPDITREKV 271  
 QY 303 NGAHSGGRDTEIHRQFGGNCDDVVSFMYLTFLEDDD-KLEQIRKDYTSGLMGTGELK 361  
 DB 272 RTHAYTGGRAVSEHRAKGGVPAEDVPFQYLSNAFFPDPAELARIEREVRAGLLSGELK 331  
 QY 362 KALIEVLQPLIAEHQARRKEVTDIIVKEMTPKLSFD 399  
 DB 332 DLAAADITEFLAHHORRAALGD--VTEALDAFRLTDD 367  
 RESULT 14  
 ID SYM METKA STANDARD; PRT; 374 AA.  
 ID Q8YTF7;  
 DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TIPBS)  
 GN TRPS OR MK0343.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OC NCBI\_TaxID=2320;  
 RN [1]\_TaxID=2320;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 RT and monophyly of archaeal methanogens."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE010331; AAM01558.1; --  
 DR HAMAP; MF\_00140; --; 1.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00579; tRNA-synt\_1b.  
 DR PRINTS; PR01039; tRNA-synt\_1b.  
 DR TIGRams; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA-TRNA\_LIGASE\_I; FALSE NEG.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 71 79 "HIGH" REGION.  
 FT SITE 247 251 "KMSKS" REGION.  
 FT SITE 374 AA; 42733 MW; 0BD455279321B828 CRC64;  
 SQ SEQUENCE 374 AA; 42733 MW; 0BD455279321B828 CRC64;  
 Query Match 21.4%; Score 453; DB 1; Length 374;  
 Best Local Similarity 32.2%; Pred. No. 2.1e-29;  
 Yatches 124; Conservative 71; Mismatches 148; Indels 42; Gaps 12;  
 QY 15 VDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQPHHFLRGIFFSHDMN 74  
 DB 2 IDPWDEE-----VDYERTEEFGRIZIDKV-----RELLPRFFPLDRGIVFGRHDYD 51  
 QY 75 QVLDAVENKPPFLYLTGRGPSSSEMHVGLIPEFTKWLQDVFNVLVIQMTDDEKVLNK 134  
 DB 52 SFLKDYNDGKLVSLGSMPSGR-MHLGHKTVVDVQLVFYQQEMDVKVYVPIADLEAHAR 110  
 QY 135 DUTLQAYGAVEN-AKCHIAAGFOIN---KTFIFSDLDYMGMSGFKVNVVKIQKHVT 189  
 DB 111 NMDLORAHRAIVAEYVLYNVAALGLDLPDRCEIYLQSE-----RKTVMALLAGRLT 164  
 QY 190 ENQVKCIFGFTSDICIGKISFPQAQPSFNSFPQIFRDRDIOCLIPCAIDQD 245  
 DB 165 KNTVKNYTGFTGTNNGHAFAPVQAADIL---HPQ-----EIEGPHRVLPVGVQDQD 215  
 QY 246 YFRMTRDVAPRIGYKPKALLHSTFFPALQCAQKMSASDPSNSIFLTDKAKIKTKV 305  
 DB 216 HLRLTEDIKAEKELIKPASTYHRTGLTGG--KMSSSKPNIAIFLTDQPFETAKKY-WN 272  
 QY 306 AFSGGRDTEIHRQFGGNCDDVVSFMYLTFLEDD--DDKLEQIRKDYTSGLMGTGELK 361

Db 273 AKTGGALEEREGGPNDECVVYELMHVHLAURIGGDEKLREIRKKCREGDIIOGEC 332  
Qy 362 KALIEVLQPLIAEHQARRKEVTDEI 386  
Db 333 RMVGEALAEILEELERREDVRDEL 357  
RESULT 15  
SYN METJA STANDARD; PRT; 370 AA.  
AC Q58810;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)  
DE (TDRS).  
GN TRPS OR MJ1415.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Sulton G.G., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Uitterback L.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
R1 "Complete genome sequence of the methanogenic archaeon, Methanococcus  
R2 jannaschii";  
R3 Science 273:1058-1073(1996).  
R4  
R5  
R6  
R7  
R8  
R9  
R10  
R11  
R12  
R13  
R14  
R15  
R16  
R17  
R18  
R19  
R20  
R21  
R22  
R23  
R24  
R25  
R26  
R27  
R28  
R29  
R30  
R31  
R32  
R33  
R34  
R35  
R36  
R37  
R38  
R39  
R40  
R41  
R42  
R43  
R44  
R45  
R46  
R47  
R48  
R49  
R50  
R51  
R52  
R53  
R54  
R55  
R56  
R57  
R58  
R59  
R60  
R61  
R62  
R63  
R64  
R65  
R66  
R67  
R68  
R69  
R70  
R71  
R72  
R73  
R74  
R75  
R76  
R77  
R78  
R79  
R80  
R81  
R82  
R83  
R84  
R85  
R86  
R87  
R88  
R89  
R90  
R91  
R92  
R93  
R94  
R95  
R96  
R97  
R98  
R99  
R100  
R101  
R102  
R103  
R104  
R105  
R106  
R107  
R108  
R109  
R110  
R111  
R112  
R113  
R114  
R115  
R116  
R117  
R118  
R119  
R120  
R121  
R122  
R123  
R124  
R125  
R126  
R127  
R128  
R129  
R130  
R131  
R132  
R133  
R134  
R135  
R136  
R137  
R138  
R139  
R140  
R141  
R142  
R143  
R144  
R145  
R146  
R147  
R148  
R149  
R150  
R151  
R152  
R153  
R154  
R155  
R156  
R157  
R158  
R159  
R160  
R161  
R162  
R163  
R164  
R165  
R166  
R167  
R168  
R169  
R170  
R171  
R172  
R173  
R174  
R175  
R176  
R177  
R178  
R179  
R180  
R181  
R182  
R183  
R184  
R185  
R186  
R187  
R188  
R189  
R190  
R191  
R192  
R193  
R194  
R195  
R196  
R197  
R198  
R199  
R200  
R201  
R202  
R203  
R204  
R205  
R206  
R207  
R208  
R209  
R210  
R211  
R212  
R213  
R214  
R215  
R216  
R217  
R218  
R219  
R220  
R221  
R222  
R223  
R224  
R225  
R226  
R227  
R228  
R229  
R230  
R231  
R232  
R233  
R234  
R235  
R236  
R237  
R238  
R239  
R240  
R241  
R242  
R243  
R244  
R245  
R246  
R247  
R248  
R249  
R250  
R251  
R252  
R253  
R254  
R255  
R256  
R257  
R258  
R259  
R260  
R261  
R262  
R263  
R264  
R265  
R266  
R267  
R268  
R269  
R270  
R271  
R272  
R273  
R274  
R275  
R276  
R277  
R278  
R279  
R280  
R281  
R282  
R283  
R284  
R285  
R286  
R287  
R288  
R289  
R290  
R291  
R292  
R293  
R294  
R295  
R296  
R297  
R298  
R299  
R300  
R301  
R302  
R303  
R304  
R305  
R306  
R307  
R308  
R309  
R310  
R311  
R312  
R313  
R314  
R315  
R316  
R317  
R318  
R319  
R320  
R321  
R322  
R323  
R324  
R325  
R326  
R327  
R328  
R329  
R330  
R331  
R332  
R333  
R334  
R335  
R336  
R337  
R338  
R339  
R340  
R341  
R342  
R343  
R344  
R345  
R346  
R347  
R348  
R349  
R350  
R351  
R352  
R353  
R354  
R355  
R356  
R357  
R358  
R359  
R360  
R361  
R362  
R363  
R364  
R365  
R366  
R367  
R368  
R369  
R370  
R371  
R372  
R373  
R374  
R375  
R376  
R377  
R378  
R379  
R380  
R381  
R382  
R383  
R384  
R385  
R386  
R387  
R388  
R389  
R390  
R391  
R392  
R393  
R394  
R395  
R396  
R397  
R398  
R399  
R400  
R401  
R402  
R403  
R404  
R405  
R406  
R407  
R408  
R409  
R410  
R411  
R412  
R413  
R414  
R415  
R416  
R417  
R418  
R419  
R420  
R421  
R422  
R423  
R424  
R425  
R426  
R427  
R428  
R429  
R430  
R431  
R432  
R433  
R434  
R435  
R436  
R437  
R438  
R439  
R440  
R441  
R442  
R443  
R444  
R445  
R446  
R447  
R448  
R449  
R450  
R451  
R452  
R453  
R454  
R455  
R456  
R457  
R458  
R459  
R460  
R461  
R462  
R463  
R464  
R465  
R466  
R467  
R468  
R469  
R470  
R471  
R472  
R473  
R474  
R475  
R476  
R477  
R478  
R479  
R480  
R481  
R482  
R483  
R484  
R485  
R486  
R487  
R488  
R489  
R490  
R491  
R492  
R493  
R494  
R495  
R496  
R497  
R498  
R499  
R500  
R501  
R502  
R503  
R504  
R505  
R506  
R507  
R508  
R509  
R510  
R511  
R512  
R513  
R514  
R515  
R516  
R517  
R518  
R519  
R520  
R521  
R522  
R523  
R524  
R525  
R526  
R527  
R528  
R529  
R530  
R531  
R532  
R533  
R534  
R535  
R536  
R537  
R538  
R539  
R540  
R541  
R542  
R543  
R544  
R545  
R546  
R547  
R548  
R549  
R550  
R551  
R552  
R553  
R554  
R555  
R556  
R557  
R558  
R559  
R560  
R561  
R562  
R563  
R564  
R565  
R566  
R567  
R568  
R569  
R570  
R571  
R572  
R573  
R574  
R575  
R576  
R577  
R578  
R579  
R580  
R581  
R582  
R583  
R584  
R585  
R586  
R587  
R588  
R589  
R590  
R591  
R592  
R593  
R594  
R595  
R596  
R597  
R598  
R599  
R600  
R601  
R602  
R603  
R604  
R605  
R606  
R607  
R608  
R609  
R610  
R611  
R612  
R613  
R614  
R615  
R616  
R617  
R618  
R619  
R620  
R621  
R622  
R623  
R624  
R625  
R626  
R627  
R628  
R629  
R630  
R631  
R632  
R633  
R634  
R635  
R636  
R637  
R638  
R639  
R640  
R641  
R642  
R643  
R644  
R645  
R646  
R647  
R648  
R649  
R650  
R651  
R652  
R653  
R654  
R655  
R656  
R657  
R658  
R659  
R660  
R661  
R662  
R663  
R664  
R665  
R666  
R667  
R668  
R669  
R670  
R671  
R672  
R673  
R674  
R675  
R676  
R677  
R678  
R679  
R680  
R681  
R682  
R683  
R684  
R685  
R686  
R687  
R688  
R689  
R690  
R691  
R692  
R693  
R694  
R695  
R696  
R697  
R698  
R699  
R700  
R701  
R702  
R703  
R704  
R705  
R706  
R707  
R708  
R709  
R710  
R711  
R712  
R713  
R714  
R715  
R716  
R717  
R718  
R719  
R720  
R721  
R722  
R723  
R724  
R725  
R726  
R727  
R728  
R729  
R730  
R731  
R732  
R733  
R734  
R735  
R736  
R737  
R738  
R739  
R740  
R741  
R742  
R743  
R744  
R745  
R746  
R747  
R748  
R749  
R750  
R751  
R752  
R753  
R754  
R755  
R756  
R757  
R758  
R759  
R760  
R761  
R762  
R763  
R764  
R765  
R766  
R767  
R768  
R769  
R770  
R771  
R772  
R773  
R774  
R775  
R776  
R777  
R778  
R779  
R780  
R781  
R782  
R783  
R784  
R785  
R786  
R787  
R788  
R789  
R790  
R791  
R792  
R793  
R794  
R795  
R796  
R797  
R798  
R799  
R800  
R801  
R802  
R803  
R804  
R805  
R806  
R807  
R808  
R809  
R810  
R811  
R812  
R813  
R814  
R815  
R816  
R817  
R818  
R819  
R820  
R821  
R822  
R823  
R824  
R825  
R826  
R827  
R828  
R829  
R830  
R831  
R832  
R833  
R834  
R835  
R836  
R837  
R838  
R839  
R840  
R841  
R842  
R843  
R844  
R845  
R846  
R847  
R848  
R849  
R850  
R851  
R852  
R853  
R854  
R855  
R856  
R857  
R858  
R859  
R860  
R861  
R862  
R863  
R864  
R865  
R866  
R867  
R868  
R869  
R870  
R871  
R872  
R873  
R874  
R875  
R876  
R877  
R878  
R879  
R880  
R881  
R882  
R883  
R884  
R885  
R886  
R887  
R888  
R889  
R890  
R891  
R892  
R893  
R894  
R895  
R896  
R897  
R898  
R899  
R900  
R901  
R902  
R903  
R904  
R905  
R906  
R907  
R908  
R909  
R910  
R911  
R912  
R913  
R914  
R915  
R916  
R917  
R918  
R919  
R920  
R921  
R922  
R923  
R924  
R925  
R926  
R927  
R928  
R929  
R930  
R931  
R932  
R933  
R934  
R935  
R936  
R937  
R938  
R939  
R940  
R941  
R942  
R943  
R944  
R945  
R946  
R947  
R948  
R949  
R950  
R951  
R952  
R953  
R954  
R955  
R956  
R957  
R958  
R959  
R960  
R961  
R962  
R963  
R964  
R965  
R966  
R967  
R968  
R969  
R970  
R971  
R972  
R973  
R974  
R975  
R976  
R977  
R978  
R979  
R980  
R981  
R982  
R983  
R984  
R985  
R986  
R987  
R988  
R989  
R990  
R991  
R992  
R993  
R994  
R995  
R996  
R997  
R998  
R999  
R1000  
R1001  
R1002  
R1003  
R1004  
R1005  
R1006  
R1007  
R1008  
R1009  
R1010  
R1011  
R1012  
R1013  
R1014  
R1015  
R1016  
R1017  
R1018  
R1019  
R1020  
R1021  
R1022  
R1023  
R1024  
R1025  
R1026  
R1027  
R1028  
R1029  
R1030  
R1031  
R1032  
R1033  
R1034  
R1035  
R1036  
R1037  
R1038  
R1039  
R1040  
R1041  
R1042  
R1043  
R1044  
R1045  
R1046  
R1047  
R1048  
R1049  
R1050  
R1051  
R1052  
R1053  
R1054  
R1055  
R1056  
R1057  
R1058  
R1059  
R1060  
R1061  
R1062  
R1063  
R1064  
R1065  
R1066  
R1067  
R1068  
R1069  
R1070  
R1071  
R1072  
R1073  
R1074  
R1075  
R1076  
R1077  
R1078  
R1079  
R1080  
R1081  
R1082  
R1083  
R1084  
R1085  
R1086  
R1087  
R1088  
R1089  
R1090  
R1091  
R1092  
R1093  
R1094  
R1095  
R1096  
R1097  
R1098  
R1099  
R1100  
R1101  
R1102  
R1103  
R1104  
R1105  
R1106  
R1107  
R1108  
R1109  
R1110  
R1111  
R1112  
R1113  
R1114  
R1115  
R1116  
R1117  
R1118  
R1119  
R1120  
R1121  
R1122  
R1123  
R1124  
R1125  
R1126  
R1127  
R1128  
R1129  
R1130  
R1131  
R1132  
R1133  
R1134  
R1135  
R1136  
R1137  
R1138  
R1139  
R1140  
R1141  
R1142  
R1143  
R1144  
R1145  
R1146  
R1147  
R1148  
R1149  
R1150  
R1151  
R1152  
R1153  
R1154  
R1155  
R1156  
R1157  
R1158  
R1159  
R1160  
R1161  
R1162  
R1163  
R1164  
R1165  
R1166  
R1167  
R1168  
R1169  
R1170  
R1171  
R1172  
R1173  
R1174  
R1175  
R1176  
R1177  
R1178  
R1179  
R1180  
R1181  
R1182  
R1183  
R1184  
R1185  
R1186  
R1187  
R1188  
R1189  
R1190  
R1191  
R1192  
R1193  
R1194  
R1195  
R1196  
R1197  
R1198  
R1199  
R1200  
R1201  
R1202  
R1203  
R1204  
R1205  
R1206  
R1207  
R1208  
R1209  
R1210  
R1211  
R1212  
R1213  
R1214  
R1215  
R1216  
R1217  
R1218  
R1219  
R1220  
R1221  
R1222  
R1223  
R1224  
R1225  
R1226  
R1227  
R1228  
R1229  
R1230  
R1231  
R1232  
R1233  
R1234  
R1235  
R1236  
R1237  
R1238  
R1239  
R1240  
R1241  
R1242  
R1243  
R1244  
R1245  
R1246  
R1247  
R1248  
R1249  
R1250  
R1251  
R1252  
R1253  
R1254  
R1255  
R1256  
R1257  
R1258  
R1259  
R1260  
R1261  
R1262  
R1263  
R1264  
R1265  
R1266  
R1267  
R1268  
R1269  
R1270  
R1271  
R1272  
R1273  
R1274  
R1275  
R1276  
R1277  
R1278  
R1279  
R1280  
R1281  
R1282  
R1283  
R1284  
R1285  
R1286  
R1287  
R1288  
R1289  
R1290  
R1291  
R1292  
R1293  
R1294  
R1295  
R1296  
R1297  
R1298  
R1299  
R1300  
R1301  
R1302  
R1303  
R1304  
R1305  
R1306  
R1307  
R1308  
R1309  
R1310  
R1311  
R1312  
R1313  
R1314  
R1315  
R1316  
R1317  
R1318  
R1319  
R1320  
R1321  
R1322  
R1323  
R1324  
R1325  
R1326  
R1327  
R1328  
R1329  
R1330  
R1331  
R1332  
R1333  
R1334  
R1335  
R1336  
R1337  
R1338  
R1339  
R1340  
R1341  
R1342  
R1343  
R1344  
R1345  
R1346  
R1347  
R1348  
R1349  
R1350  
R1351  
R1352  
R1353  
R1354  
R1355  
R1356  
R1357  
R1358  
R1359  
R1360  
R1361  
R1362  
R1363  
R1364  
R1365  
R1366  
R1367  
R1368  
R1369  
R1370  
R1371  
R1372  
R1373  
R1374  
R1375  
R1376  
R1377  
R1378  
R1379  
R1380  
R1381  
R1382  
R1383  
R1384  
R1385  
R1386  
R1387  
R1388  
R1389  
R1390  
R1391  
R1392  
R1393  
R1394  
R1395  
R1396  
R1397  
R1398  
R1399  
R1400  
R1401  
R1402  
R1403  
R1404  
R1405  
R1406  
R1407  
R1408  
R1409  
R1410  
R1411  
R1412  
R1413  
R1414  
R1415  
R1416  
R1417  
R1418  
R1419  
R1420  
R1421  
R1422  
R1423  
R1424  
R1425  
R1426  
R1427  
R1428  
R1429  
R1430  
R1431  
R1432  
R1433  
R1434  
R1435  
R1436  
R1437  
R1438  
R1439  
R1440  
R1441  
R1442  
R1443  
R1444  
R1445  
R1446  
R1447  
R1448  
R1449  
R1450  
R1451  
R1452  
R1453  
R1454  
R1455  
R1456  
R1457  
R1458  
R1459  
R1460  
R1461  
R1462  
R1463  
R1464  
R1465  
R1466  
R1467  
R1468  
R1469  
R1470  
R1471  
R1472  
R1473  
R1474  
R1475  
R1476  
R1477  
R1478  
R1479  
R1480  
R1481  
R1482  
R1483  
R1484  
R1485  
R1486  
R1487  
R1488  
R1489  
R1490  
R1491  
R1492  
R1493  
R1494  
R1495  
R1496  
R1497  
R1498  
R1499  
R1500  
R1501  
R1502  
R1503  
R1504  
R1505  
R1506  
R1507  
R1508  
R1509  
R1510  
R1511  
R1512  
R1513  
R1514  
R1515  
R1516  
R1517  
R1518  
R1519  
R1520  
R1521  
R1522  
R1523  
R1524  
R1525  
R1526  
R1527  
R1528  
R1529  
R1530  
R1531  
R1532  
R1533  
R1534  
R1535  
R1536  
R1537  
R1538  
R1539  
R1540  
R1541  
R1542  
R1543  
R1544  
R1545  
R1546  
R1547  
R1548  
R1549  
R1550  
R1551  
R1552  
R1553  
R1554  
R1555  
R1556  
R1557  
R1558  
R1559  
R1560  
R1561  
R1562  
R1563  
R1564  
R1565  
R1566  
R1567  
R1568  
R1569  
R1570  
R1571  
R1572  
R1573  
R1574  
R1575  
R1576  
R1577  
R1578  
R1579  
R1580  
R1581  
R1582  
R1583  
R1584  
R1585  
R1586  
R1587  
R1588  
R1589  
R1590  
R1591  
R1592  
R1593  
R1594  
R1595  
R1596  
R1597  
R1598  
R1599  
R1600  
R1601  
R1602  
R1603  
R1604  
R1605  
R1606  
R1607  
R1608  
R1609  
R1610  
R1611  
R1612  
R1613  
R1614  
R1615  
R1616  
R1617  
R1618  
R1619  
R1620  
R1621  
R1622  
R1623  
R1624  
R1625  
R1626  
R1627  
R1628  
R1629  
R1630  
R1631  
R1632  
R1633  
R1634  
R1635  
R1636  
R1637  
R1638  
R1639  
R1640  
R1641  
R1642  
R1643  
R1644  
R1645  
R1646  
R1647  
R1648  
R1649  
R1650  
R1651  
R1652  
R1653  
R1654  
R1655  
R1656  
R1657  
R1658  
R1659  
R1660  
R1661  
R1662  
R1663  
R1664  
R1665  
R1666  
R1667  
R1668  
R1669  
R1670  
R1671  
R1672  
R1673  
R1674  
R1675  
R1676  
R1677  
R1678  
R1679  
R1680  
R1681  
R1682  
R1683  
R1684  
R1685  
R1686  
R1687  
R1688  
R1689  
R1690  
R1691  
R1692  
R1693  
R1694  
R1695  
R1696  
R1697  
R1698  
R1699  
R1700  
R1701  
R1702  
R1703  
R1704  
R1705  
R1706  
R1707  
R1708  
R1709  
R1710  
R1711  
R1712  
R1713  
R1714  
R1715  
R1716  
R1717  
R1718  
R1719  
R1720  
R1721  
R1722  
R1723  
R1724  
R1725  
R1726  
R1727  
R1728  
R1729  
R1730  
R1731  
R1732  
R1733  
R1734  
R1735  
R1736  
R1737  
R1738  
R1739  
R1740  
R1741  
R1742  
R1743  
R1744  
R1745  
R1746  
R1747  
R1748  
R1749  
R1750  
R1751  
R1752  
R1753  
R1754  
R1755  
R1756  
R1757  
R1758  
R1759  
R1760  
R1761  
R1762  
R1763  
R1764  
R1765  
R1766  
R1767  
R1768  
R1769  
R1770  
R1771  
R1772  
R1773  
R1774  
R1775  
R1776  
R1777  
R1778  
R1779  
R1780  
R1781  
R1782  
R1783  
R1784  
R1785  
R1786  
R1787  
R1788  
R1789  
R1790  
R1791  
R1792  
R1793  
R1794  
R1795  
R1796  
R1797  
R1798  
R1799  
R1800  
R1801  
R1802  
R1803  
R1804  
R1805  
R1806  
R1807  
R1808  
R1809  
R1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:42:41 ; Search time 117 Seconds  
(without alignments)  
1081.391 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNEGPDATEBDFVDPWTV.....VTDEIVKEFMPKLSDFQ 401

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTEMBL 25:\*

1: sp.archaea:\*

2: sp.bacteria:\*

3: sp.fungi:\*

4: sp.human:\*

5: sp.invertebrate:\*

6: sp.mammal:\*

7: sp.mhc:\*

8: sp.organelle:\*

9: sp.phage:\*

10: sp.plant:\*

11: sp.rodent:\*

12: sp.virus:\*

13: sp.vertibrate:\*

14: sp.unclassified:\*

15: sp.virus:\*

16: sp.bacteriap:\*

17: sp.archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1957	92.5	475	11 Q9DC65	Q9dc65 mus musculus
2	1957	92.5	481	11 Q99J58	Q99j58 mus musculus
3	1953	92.3	475	11 Q8CZY4	Q8czy4 mus musculus
4	1820	86.0	475	13 Q7ZWT7	Q7zwt7 xenopus lae
5	1537	72.6	305	11 Q70184	Q70184 cavia porce
6	1374.5	65.0	420	5 Q9U4Y0	Q9u4y0 drosophila
7	1374.5	65.0	430	5 Q9U4Y1	Q9u4y1 drosophila
8	1370.5	64.8	430	5 Q9VHG2	Q9vhg2 drosophila
9	1323	62.5	417	5 Q9VLR2	Q9vlr2 caenorhabdi
10	1301	61.5	402	10 Q9SKL5	Q9skl5 arabidopsis
11	1209	57.1	491	3 Q8700C	Q8700c neurospora
12	1026	48.5	632	5 Q8LDW3	Q8ldw3 plasmodium
13	942	44.5	385	5 Q8SQY5	Q8sqy5 encephalito
14	795.5	37.6	490	5 Q9U1F5	Q9u1f5 leishmania
15	682	32.2	136	6 Q95295	Q95295 sus scrofa
16	400.5	18.9	165	10 Q7XY51	Q7xy51 griffithsia

17	310	14.7	111	5 Q95YL8	Q95yl8 encephalito
18	296.5	14.0	136	6 Q9TS88	Q9ts88 bos taurus
19	294	13.9	157	5 Q9U533	Q9u533 trypanosoma
20	286	13.5	109	4 Q9UD15	Q9ud15 homo sapien
21	289.5	12.7	513	17 Q9HN83	Q9hn83 halobacteri
22	203.5	9.6	341	16 Q8E2J5	Q8e2j5 streptococc
23	203.5	9.6	341	16 Q8DWP7	Q8dwp7 streptococc
24	191	9.0	324	17 Q8TXZ2	Q8txz2 methanopyru
25	188	8.9	364	17 Q9YA64	Q9ya64 aeropyrum p
26	185.5	8.8	366	16 Q8ALC7	Q8al07 bacteroides
27	181.5	8.6	331	17 Q975Z1	Q97921 thermoplasm
28	181	8.6	340	16 Q8DRR1	Q8drr1 streptococc
29	178	8.4	340	16 Q87Q03	Q87qg3 vibrio para
30	174	8.2	334	16 Q832J8	Q832j8 enterococcu
31	171.5	8.1	351	10 Q93018	Q93018 arabidopsis
32	171.5	8.1	385	10 Q8S9T2	Q8s9t2 arabidopsis
33	170.5	8.1	376	5 Q7YAA0	Q7yya0 cryptospori
34	167	7.9	351	16 Q9RVD6	Q9rvd6 deinococcus
35	166.5	7.9	338	16 Q8BZB4	Q8bz24 lactobacill
36	164.5	7.8	377	5 Q86A90	Q86a90 dictyosteli
37	164.5	7.8	895	10 Q9SGN2	Q9sgn2 arabidopsis
38	158.5	7.5	356	16 Q83A61	Q83a61 coxiella bu
39	158.5	7.5	375	17 Q9VQ27	Q9v027 pyrococcus
40	156	7.4	372	17 Q8Z1F7	Q8zy07 pyrobaculum
41	154.5	7.3	102	1 Q07119	Q07119 halobacteri
42	153.5	7.3	317	17 Q8TS11	Q8ts11 methanosaer
43	152.5	7.2	682	5 Q9N9B8	Q9n9b8 leishmania
44	152	7.2	317	17 Q8PVK0	Q8pvk0 methanosaer
45	149.5	7.1	327	17 Q9HN62	Q9hn62 halobacteri

## ALIGNMENTS

### RESULT 1

Q9DC65 Q9DC65 PRELIMINARY; PRT; 475 AA.  
AC Q9DC65;  
DT 01-JUN-2001 (TRMBLrel. 17, Created)  
DT 01-JUN-2001 (TRMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)  
DE Adult male lung cDNA, RIKEN full-length enriched library,  
clone:1200002C07, full insert sequence.  
GN WARS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Lung;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli C., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RT Hayashizaki Y.  
RL "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
EMBL; AK004542; BAB23357.1; -.

DR MGD; MGI:104630; Wars.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; tTP tRNA-synt\_1b.  
 DR InterPro; IPR000738; WHEP-TRS.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; tRNA-synt\_1b.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS00178; AA tRNA LIGASE\_I; 1.  
 DR PROSITE; PS00742; WHEP-TRS; 1.  
 DR PROSITE; PS00742; WHEP-TRS; 1.  
 DR Aminoacyl-tRNA synthetase.  
 DR KW Aminoacyl-tRNA synthetase.  
 DR SEQUENCE 475 AA; 53641 MW; C3467FE95521DB4C CRC64;  
 Query Match 92.5%; Score 1957; DB 11; Length 475;  
 Best Local Similarity 91.5%; Pred. No. 6.7e-163;  
 Matches 366; Conservative 20; Mismatches 14; Indels 0; Gaps 0;  
 QY 2 NHGPDATAEEDFVDPWTVTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHF 61  
 DI 76 NCDSDATKASEDFVDPWTVTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHF 135  
 QY 62 LRGGIFFSHRDMQVLDAYENKKPFYLYTGRGSPSEAMHVGHLIPFTTKWLQDVFNVP 121  
 DI 136 LRGGIFFSHRDMQVLDAYENKKPFYLYTGRGSPSEAMHVGHLIPFTTKWLQDVFNVP 195  
 QY 122 VIQMTDDEKYLKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFGYKNV 181  
 DI 196 VIQMSDDEKYLKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFGYKNV 255  
 QY 182 VKIQKHTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAI 241  
 DI 256 VKIQKHTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAI 315  
 QY 242 DQDYPFRMTDVAAPRIGYKPKALLHSTFFPALQAPSPNSFPQIFRDRDIOCLIPCAI 301  
 DI 316 DQDYPFRMTDVAAPRIGYKPKALLHSTFFPALQAPSPNSFPQIFRDRDIOCLIPCAI 375  
 QY 302 VNKHAFSGGRTTIEHRQFGNCEDVDVSPMYLTFLEDDDKLEQIRKDYTSGLMTGELK 361  
 DI 376 VNKHAFSGGRTTIEHRQFGNCEDVDVSPMYLTFLEDDDKLEQIRKDYTSGLMTGELK 435  
 QY 362 KALIEVLQPLIAEHQARRKEVTDEIVKEPMTPKLSDFQ 401  
 DI 436 KTLIDVLQPLIAEHQARRKAVTETVKEPMTPKLSDFQ 475

## RESULT 2

Q95J58  
 ID Q95J58 PRELIMINARY; PRT; 481 AA.  
 AC Q95J58;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase.  
 GN WARS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC003450; AAH03450.1;  
 DR MGD; MGI:104630; Wars.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt\_1b.

DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; tTP tRNA-synt\_1b.  
 DR InterPro; IPR000738; WHEP-TRS.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR PRINTS; PR01039; tRNA-synt\_1b.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS00178; AA tRNA LIGASE\_I; 1.  
 DR PROSITE; PS00742; WHEP-TRS; 1.  
 DR Aminoacyl-tRNA synthetase.  
 DR KW Aminoacyl-tRNA synthetase.  
 DR SEQUENCE 481 AA; 54325 MW; A754ELDDP58E2EF3 CRC64;

Query Match 92.5%; Score 1957; DB 11; Length 481;  
 Best Local Similarity 91.5%; Pred. No. 6.8e-163;  
 Matches 366; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

QY 2 NHGPDATAEEDFVDPWTVTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHF 61  
 DI 76 NCDSDATKASEDFVDPWTVTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHF 135  
 QY 62 LRGGIFFSHRDMQVLDAYENKKPFYLYTGRGSPSEAMHVGHLIPFTTKWLQDVFNVP 121  
 DI 136 LRGGIFFSHRDMQVLDAYENKKPFYLYTGRGSPSEAMHVGHLIPFTTKWLQDVFNVP 195  
 QY 122 VIQMTDDEKYLKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFGYKNV 181  
 DI 196 VIQMSDDEKYLKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFGYKNV 255  
 QY 182 VKIQKHTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAI 241  
 DI 256 VKIQKHTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAI 315  
 QY 242 DQDYPFRMTDVAAPRIGYKPKALLHSTFFPALQAPSPNSFPQIFRDRDIOCLIPCAI 301  
 DI 316 DQDYPFRMTDVAAPRIGYKPKALLHSTFFPALQAPSPNSFPQIFRDRDIOCLIPCAI 375  
 QY 302 VNKHAFSGGRTTIEHRQFGNCEDVDVSPMYLTFLEDDDKLEQIRKDYTSGLMTGELK 361  
 DI 376 VNKHAFSGGRTTIEHRQFGNCEDVDVSPMYLTFLEDDDKLEQIRKDYTSGLMTGELK 435  
 QY 362 KALIEVLQPLIAEHQARRKEVTDEIVKEPMTPKLSDFQ 401  
 DI 436 KTLIDVLQPLIAEHQARRKAVTETVKEPMTPKLSDFQ 475

## RESULT 3

Q80ZY4  
 ID Q80ZY4 PRELIMINARY; PRT; 475 AA.  
 AC Q80ZY4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046232; AAH46232.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; tTP tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR PRINTS; PR01039; tRNA-synt\_1b.

13 TIGRFAMs; TIGR00233; tips; 1.  
 14 PROSITE; PS00178; AA TRNA\_LIGASE\_I; 1.  
 15 PROSITE; PS00762; WHEP\_TRS; 1.  
 16 Aminoacyl-tRNA synthetase.  
 17 SEQUENCE 475 AA; 53613 MW; D841D8B26973F214 CRC64;  
 Query Match 92.3%; Score 1953; DB 11; Length 475;  
 Best Local Similarity 91.2%; Pred. No. 1.5e-162;  
 Matches 365; Conservative 20; Mismatches 15; Indels 0; Gaps 0;  
 2 NHGPDATAEEDFVDPWTQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHE 61  
 76 NCSDATKASEDFVDPWTARTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHE 135  
 62 LRRGIFFSHRDMQVLDAYENKPPFLYTGROPSSEAMHVGHLIPFIPTKWLQDVFNVP 121  
 136 LRRGIFFSHRDMQVLDAYENKPPFLYTGROPSSEAMHVGHLIPFIPTKWLQDVFNVP 195  
 122 VIQMTDDKYLWKQJLDOAYDAVENAKDIIACGFDINKTFISDLDYMGSSGFGYKNV 181  
 196 VIQMTDDKYLWKQJLDOAYDAVENAKDIIACGFDINKTFISDLDYMGSSGFGYKNV 255  
 182 VKIQKHVTNFVQKGFEGTDSICIGKISFPALQAAAPSFNSPQIFRDRDTIQCLIPCAI 241  
 256 VKIQKHVTNFVQKGFEGTDSICIGKISFPALQAAAPSFNSPQIFRDRDTIQCLIPCAI 315  
 242 DQDPYFRMTRDVAPRIGVPKPKALLHSTFPALQGAQTKVKSADPNSSIFLTDTAKQIKT 301  
 316 DQDPYFRMTRDVAPRIGVPKPKALLHSTFPALQGAQTKVKSADPNSSIFLTDTAKQIKS 375  
 302 VNKHAFSGGRDTIEHRQFGGNCVDVDFVSMYLTFFLEDDKLEQIRKDYTSGMLTGEUK 361  
 376 VNKHAFSGGRDTIEHRQFGGNCVDVDFVSMYLTFFLEDDKLEQIRKDYTSGMLTGEUK 435  
 362 KALIEVLQPLIAEHQARRKEVTEIVKEMTPRKLSDFQ 401  
 436 KTLIDVLQPLIAEHQARRKAVTEETVKEMTPRKLSDFHQ 475  
 PRELIMINARY; PRT; 475 AA.  
 RESULT 4  
 Q72WT7  
 ID Q72WT7 PRELIMINARY; PRT; 475 AA.  
 AC Q72WT7;  
 DT 01-JUN-2003 (TIGR00233; tips; 1.  
 DT 01-JUN-2003 (TIGR00233; tips; 1.  
 DT 01-OCT-2003 (TIGR00233; tips; 1.  
 DE Similar to tryptophanyl-tRNA synthetase.  
 OS Xenopus laevis (African clawed frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 CC Xenopodinae; Xenopus.  
 CX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S.; Strausberg R.;  
 RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046713; AAH46713.1; -;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; trna-synt\_1b.  
 DR InterPro; IPR002306; trna-synt\_1b.  
 DR InterPro; IPR002307; trna-synt\_1b.  
 DR InterPro; IPR002308; trna-synt\_1b.  
 DR Pfam; PF00579; trna-synt\_1b; 1.  
 DR Pfam; PF00580; WHEP-TRS; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; tips; 1.  
 DR PROSITE; PS00178; AA TRNA\_LIGASE\_I; 1.  
 DR PROSITE; PS00762; WHEP\_TRS; 1.  
 DR Aminoacyl-tRNA synthetase.  
 SEQUENCE 475 AA; 53830 MW; CDCCE6DB08AE357 CRC64;

Query Match 86.0%; Score 1820; DB 13; Length 475;  
 Best Local Similarity 83.2%; Pred. No. 6.9e-151;  
 Matches 333; Conservative 39; Mismatches 28; Indels 0; Gaps 0;  
 1 SNHGPDATAEEDFVDPWTQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHE 60  
 76 NRGPSPTNDGDDFVDPWTQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHE 135  
 61 FLRRGIFFSHRDMQVLDAYENKPPFLYTGROPSSEAMHVGHLIPFIPTKWLQDVFNVP 120  
 136 FLRRGIFFSHRDMQVLDAYENKPPFLYTGROPSSEAMHVGHLIPFIPTKWLQDVFNVP 195  
 121 LVIQMTDDKYLWKQJLDOAYDAVENAKDIIACGFDINKTFISDLDYMGSSGFGYKN 180  
 196 LVVQLTDDKYLWKQJLDOAYDAVENAKDIIACGFDINKTFISDLDYMGSSGFGYKN 255  
 181 VKIQKHVTNFVQKGFEGTDSICIGKISFPALQAAAPSFNSPQIFRDRDTIQCLIPCA 240  
 256 VKIQKHVTNFVQKGFEGTDSICIGKISFPALQAAAPSFNSPQIFRDRDTIQCLIPCA 315  
 241 IQDPYFRMTRDVAPRIGVPKPKALLHSTFPALQGAQTKVKSADPNSSIFLTDTAKQIKT 300  
 316 IQDPYFRMTRDVAPRIGVPKPKALLHSTFPALQGAQTKVKSADPNSSIFLTDTAKQIKS 375  
 301 VNKHAFSGGRDTIEHRQFGGNCVDVDFVSMYLTFFLEDDKLEQIRKDYTSGMLTGEUK 360  
 376 VNKHAFSGGRDTIEHRQFGGNCVDVDFVSMYLTFFLEDDKLEQIRKDYTSGMLTGEUK 435  
 361 KALIEVLQPLIAEHQARRKEVTEIVKEMTPRKLSDFQ 400  
 436 KKILTETLQPLIAEHQARRKHTEETVKQFMMPKRLAEDF 475  
 PRELIMINARY; PRT; 305 AA.  
 RESULT 5  
 Q72184  
 ID Q72184 PRELIMINARY; PRT; 305 AA.  
 AC Q72184;  
 DT 01-AUG-1998 (TIGR00233; tips; 1.  
 DT 01-AUG-1998 (TIGR00233; tips; 1.  
 DT 01-OCT-2003 (TIGR00233; tips; 1.  
 DE Tryptophan-tRNA synthetase (Fragment).  
 CS Cavia porcellus (Guinea pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 CX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley; TISSUE=Spleen;  
 RA Yang D.; Goto R.; Watanabe N.; Kobayashi Y.;  
 RA Identification and Cloning of Genes Whose Expressions are Elevated  
 RT during DNCB-induced Guinea Pig Skin Delayed-type Hypersensitivity  
 RT Reaction.  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A501222; BAA25288.1; -;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; trna-synt\_1b.  
 DR InterPro; IPR002306; trna-synt\_1b.  
 DR Pfam; PF00579; trna-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 KW Aminoacyl-tRNA synthetase.  
 FT NON TER 1  
 SQ SEQUENCE 305 AA; 34838 MW; 093764333780EB7 CRC64;  
 Query Match 72.6%; Score 1537; DB 11; Length 305;  
 Best Local Similarity 94.4%; Pred. No. 2.4e-126;  
 Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;  
 97 EAMHVGHLIPFIPTKWLQDVFNVPVIVOMTDEKYLWKQJLDOAYDAVENAKDIIACG 156  
 1 EAMHVGHLIPFIPTKWLQDVFNVPVIVOMTDEKYLWKQJLDOAYDAVENAKDIIACG 60

Q	157	FDINKTIFSDLDYMGSSGFYKVVVKIQKHVTENOVKGI	FGFTSDSDCIGKTSPPAIQAA	216
D	61	FDINKTIFSDLEYMGSEGFYKVVVKIQKHVTENOVKGI	FGFTSDSDCIGKLSPPAVQAA	120
Q	217	PSFNSFPQIQRDRDIOCLIPCAIDQDPYFMTDRDVAPRIGYPKPALJHSTFFPALQGA	276	
D	121	PSFNSFPQIQRDRDIOCLIPCAIDQDPYFMTDRDVAPRIGYPKPALJHSTFFPALQGA	180	
Q	277	QTKASASDPNSSIFLTDYAKQIKTKVNHAFSGGRDITTEHROFGCNCDDVYSFWMLTFF	336	
D	181	QTKASASDPNSSIFLTDYAKQIKTKVNHAFSGGRDITTEHROFGCNCDDVYSFWMLTFF	240	
Q	337	LEDDDKLEQIRKDYTSYGAMLTGELKALIEVLQPLIAEHQARRKEVTDIEIVKEFMTPRKL	396	
D	241	LEDDDKLEQIRKDYTSYGAMLTGELKALIEVLQPLIAEHQARRKEVTDIEIVKEFMTPRKL	300	
Q	397	SDFEQ	401	
D	301	SFHFO	305	

```

RESULT 6
Q9U4Y0 PRELIMINARY; PRT; 420 AA.
ID AC
C Q9U4Y0;
DT 01-MAY-2000 (TrEMBLrel_13, Created)
DT 01-MAY-2000 (TrEMBLrel_13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel_25, Last annotation update)
DE TricryptophanyI-tRNA synthetase (Fragment).
OS AATs-TRP OR CG9735.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=9250164; PubMed=10233165;
RX Sessaiah P., Andrew D.J.;
RA "WR8-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
RT the developing Drosophila salivary gland.";
RL Mol. Biol. Cell 10:1595-1608(1999).
RE EMBL; AF125157; AAF20167.1; -.
DR FlyBase; FBgn0010803; Aats-trp.
DR GO; GO:0005528; P:ATP binding; IEA.
DR GO; GO:0004830; P:tryptophan-CRNA ligase activity; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; trna-synt_1b.
DR InterPro; IPR01412; trna-synt_1.
DR InterPro; IPR002306; Ttp trna-synt_lb.
DF Pfam; PF00579; tRNA-synt_lb; -.
DF PRINTS; PR01039; TRNASYNTHTRP.
DF TIGRFAMs; TIGR00233; trps; 1.
DF PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KV Aminoacyl-tRNA synthetase.
ON NON TER
RN 1
SEQUENCE 420 AA: 45675 MW: 447000.00

```

[illegible]

148	Db	DDEKTLWKDLKVEDDAIKLGRNAKDIVAIGFDVNKTFIFNNLEFVGKCPAMYQNIIRIQK	207
187	Qy	HVTFNVOVKGIFGFTDSOCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCAIDQDQY	246
208	Db	CVTENNVOVKGIFGFGDSDIIGKIGFPAQAAPALSTFFPIFGNR-KVHCLIPCAIDQDQY	266
247	Qy	FRMTRDVAPRLGVKPPALLHSTFPPALQGAQTKMSASDPNSISFILTDAKQIK-KVNKHA	306
267	Db	FRMTRDVAPRLGVKPPKALIHSTFPPALQGAQTKMSASDPNSAVILTDTPKQIKNKINKYA	326
307	Qy	FSGGRDTRIEHRFGGNCDDVDSFMVLTFFLEDDDKLEIQIEKDYTSGAWLTGELKALIE	366
327	Db	FSGGRVTVBHRKLGGVVEVYVSQYLLKFFLEDDAKLEIEVTVAYKSGEMLTGEIKKLAVE	386
367	Qy	VLOPLTAHQARKEVTDDEIVKGFMTPRKLSF	398
387	Db	TLTPVYEQHOARAKLITDEVLDKYYFELRPLKF	418

RESULT 7

Q904Y1

ID Q904Y1 PRELIMINARY; PRT; 430 AA.

DC Q904Y1,

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Tryptophanyl-tRNA synthetase (AT21437p).

GN AATS-TRP OR CG9735.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Hexapoda; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

CC NCBI\_TaxID=7227;

LN [1]

LN [2]

SEQUENCE FROM N.A.

MEDLINE=99250164; PubMed=10233165;

Seshiah P., Andrew D.J.;

"WIS-95D: A tryptophanyl-tRNA synthetase expressed to high levels in the developing Drosophila salivary gland.";

Mol. Biol. Cell 10:1595-1608(1999).

SEQUENCE FROM N.A.

Scapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G., Miranda A., Murgall C.J., Nunco J., Pacleb J., Paragas V., Park S., Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celniker S.;

Submitted [JAN-2002] to the EMBL/GenBank/DBJ databases.

EMBL; AF125156; AAF20166.1; --

EMBL; AY075249; AAL68116.1; --

FlyBase; FBgn0010803; Aats-trp.

GO; GO:0005524; F-ATP binding; IEA.

GO; GO:0004830; F-tryptophan-tRNA ligase activity; IEA.

GO; GO:0006436; P-tryptophanyl-tRNA aminoacylation; IEA.

InterPro; IPR023055; tRNA-synt lb.

InterPro; IPR001412; tRNA-synt I.

InterPro; IPR002306; Trp tRNA-synt lb.

Pfam; PRF00573; tRNA-synt lb; 1.

PRINTS; PR01039; TRNASYNTHTRP.

TIGRFAMS; TIGR00233; tids; 1.

PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.

Aminoacyl-tRNA synthetase.

SEQUENCE 430 AA; 47995 MW; 2E3F9E9BC1E9979A CRC64;

Query Match 65.0%; Score 1374.5; DB 5; Length 430;  
Best Local Similarity 64.8%; Pred. No. 7e-112;  
Matches 234; Conservative 61; Mismatches 76; Indels 1; Gaps 1;



QY 67 FFSHRDNNQVLDAYENKKPFVLYTGRGSSSEAMVGHLPFIPTFKWLODVFNVLVIOMT 126  
 DB 98 FFSHRDLHTLTLEQGGKPFVLYTGRGSSSLVGHLPFIMTKWLOQETEDVPLVIQIT 157  
 QY 127 DDEKYLWKOLTLDOAYGDVAENAKDIICAGFDINKTIFPSDLDYMGSSGGYKKNVVKIQ 186  
 DB 158 DDEKTLWKOLKVEDAIKLGRENKDIIVAIGDVKNTIFNNLEFVKGKPMYQMIIRIQ 217  
 QY 187 HVTFNQVKGIFGTDSICIKISPAIOAPSFNSPFIQIFRDRDIOCLIPCAIDQDPY 246  
 DB 218 CVTFNQVKGIFGFGDSIIGKIGPAAQAPASITSTPFIFGNE-KVHCLIPCAIDQDPY 276  
 QY 247 FRMTDVAIPRIGVXPALLHSTFPPALOGATKYSASDPNSSIFLTATKQIKTKVAKHA 306  
 DB 277 FRMTDVAIPRIGVXPALLHSTFPPALOGATKYSASDPNSSIFLTATKQIKTKVAKHA 336  
 QY 307 FSGGRDTIEHROFGGNCDDVDVFMVLTFFLEDECKLEQIRKDYTSGLMTLGLKXALIE 366  
 DB 337 FSGGRVTEHRRKLGGVPEVDVSYQL-KFLEDDAKLEEVVAYSGEMLTGEIKKLAVE 396  
 QY 367 VLOPLIAHQARKEVTDVIEKFMTPRKUSP 398  
 DB 397 TLTPIVEQHQARKLTIDEVLDKYFELEPLKF 428

RESULT 8  
 QSVHG2  
 ID QSVHG2 PRELIMINARY; PRT; 430 AA.  
 AC QSVHG2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG9735-PA.  
 GN AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RXN [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazee R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch K., Baldwin D.,  
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Surtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cterry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.H., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,  
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel J., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson S.D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese Y.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J.J., Baldwin J., Bonzon J., Beeson K.V., Busan D.A.,  
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Douc L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwan C., Jallali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
 RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler J.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris K.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RA FlyBase;  
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AB03682; AAF54352.; -;  
 DR FlyBase; FBgn0010803; Aacs-trp.  
 DR GO; GO:0005524; F-ATP binding; IEA.  
 DR GO; GO:0006436; P-tryptophanyl-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P-tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1.  
 DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 DR SEQUENCE 430 AA; 47971 MW; 2938EECC69E979F CRC64;

Query Match 64.8%; Score 1370.5; DB 5; Length 430;  
 Best Local Similarity 64.5%; Pred. No. 1.6e-111;  
 Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;  
 QY 7 ATEAEEDFVDPTVTQTSAGKIDYDKLIVRFSGSKIDKELINRIERATGQRPHHFLRGI 66  
 DB 38 ATAPTEDVDVDFNVASSNDAGVDYDKLIRFGSSKIDBELIARFEKITGKPAHHFIRRG 97  
 QY 67 FFSHRDNNQVLDAYENKKPFVLYTGRGSSSEAMVGHLPFIPTFKWLODVFNVLVIOMT 126  
 DB 98 FFSHRDLHTLTLEQGGKPFVLYTGRGSSSLVGHLPFIMTKWLOQETEDVPLVIQIT 157  
 QY 127 DDEKYLWKOLTLDOAYGDVAENAKDIICAGFDINKTIFPSDLDYMGSSGGYKKNVVKIQ 186  
 DB 158 DDEKTLWKOLKVEDAIKLGRENKDIIVAIGDVKNTIFNNLEFVKGKPMYQMIIRIQ 217

```

2Y 187 HVTFNQVKGIFGFTSDSDICIGKISFPALQAAPSPNSFPQIFRDRDITDQCLIPCAIDQDPY 246
2b 218 CVTFNQVKGIFGFTSDSDIIGKIGFPAAQAAPASSTPFIFGNR-KVHCLIPCAIDQDPY 276
2Y 247 FMTROVAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKXHA 306
2b 277 FMTROVAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKXHA 336
2Y 307 FSGGRDTIEHROFGGNCVDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 366
2b 337 FSGGRVSVVEHRKLGVPDVVSQLLKFLFLEDDAKLEEVVAVSKGEMLTGHIKLAIVE 396
2Y 367 VLOPLIAEHOARKEVTDIVKPFMTPRKLSF 398
2b 397 TLTPIVEHQOARKEVTDIVKPFMTPRKLSF 428

RESULT 9
19ULR2 PRELIMINARY; PRT; 417 AA.
Q9ULR2
01-MAY-2000 (TremBLrel. 13, Created)
01-OCT-2003 (TremBLrel. 25, Last sequence update)
01-OCT-2003 (TremBLrel. 25, Last annotation update)
C. elegans WRS-1 protein (Corresponding sequence Y8CD3A.1)
WRS-1.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]
SEQUENCE FROM N.A.
STRAIN=Bristol N2;
MEDLINE=93069613; PubMed=9851916;
note;
"Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
[2]
Science 282:2012-2018(1998).
[3]
SEQUENCE FROM N.A.
STRAIN=Bristol N2;
Harris B.R.;
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
Ohtsuki T.;
"C. elegans tryptophanyl-tRNA synthetases.";
Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
Kershaw J.K.;
Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
EMBL; AL021470; CAB60439.2; JOINED.
EMBL; AL132853; CAB60439.2; -
EMBL; AB096635; BAC76735.1; -
EMBL; AL021470; CAB17953.1; -
EMBL; AL132853; CAB17953.1; JOINED.
Aminoacyl-tRNA synthetase.
SEQUENCE 417 AA; 47238 MW; 7717FE4F578BD0A CRC64;

Query Match 62.5%; Score 1323; DB 5; Length 417;
Best Local Similarity 62.0%; Pred. No. 2.2e-107;
Matches 246; Conservative 63; Mismatches 86; Indels 2; Gaps 1;

4 GPDAREAEVDPWTVQTSAGKIDYDKLIVRSGSKIDKELINLTERATGQPHFIR 63
23 GGGVQDEEDRVTFWETTTKATGIDKILVFGCKLDEIIARFVTVGHKASPMR 82
64 RGTFHGRDMQVLDAYENKKPFYLTGRCPSAMHGHILIPFTKWLQDVNFVLVI 123
83 RGMFFAHRDLTALDRKEQKPFYLTGRCASSGLHGLVFPFTKWLQDVNFVLVI 142

```

```

QY 124 QMTDDEKYLKOLTLDOAYDAVENAKDIACGPDINKTFISDLDMGSSGYKNVVK 183
Db 143 QMTDDEKFLKMKVDKAEKQARENKDIISVGFDPKTKTFIANNFY--KCPPEYENIVK 200
QY 184 IQKHVTFNQVKGIFGFTSDSDICIGKISFPALQAAPSPNSFPQIFRDRDITDQCLIPCAIDQ 243
Db 201 IKWVNTQARALFGTTPEDCLGAAPFAVEAAPCFASSFPQIFGRNDIPCLIPCAIDQ 260
QY 244 DPFMTROVAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKX 303
Db 261 DPFMTROVAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKX 320
QY 304 KHAFFSGGRDTIEHROFGGNCVDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 363
Db 321 KYAFSGGQOTVQSHRKGKGGNCVDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 380
QY 364 LIEVLQPLIAEHOARKEVTDIVKPFMTPRKLSF 400
Db 381 ATQKVCIEVLEMQERKLVTDIVKPFMTPRKLSF 417

RESULT 10
Q9SR15 PRELIMINARY; PRT; 402 AA.
AC Q9SR15
01-MAY-2000 (TremBLrel. 13, Created)
01-MAY-2000 (TremBLrel. 13, Last sequence update)
01-OCT-2003 (TremBLrel. 25, Last annotation update)
Putative tryptophanyl-tRNA synthetase.
GN F7018.7 OR AT3G04600.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
Lin X.; Kaul S.; Town C.D.; Benito X.; Creasy T.H.; Haas B.;
Ronning C.M.; Koo H.; Fujii C.Y.; Utterback T.R.; Barnstead M.E.;
Bowman C.L.; White O.; Nierman W.C.; Fraser C.M.;
"Arabidopsis thaliana chromosome III BAC F7018 genomic sequence.";
Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
Yamada K.; Liu S.X.; Sakano H.; Pham P.K.; Banh J.; Chung M.K.;
Goldsmith A.D.; Lee J.M.; Quach H.L.; Toriumi M.; Yu G.; Bowser L.;
Carninci P.; Chen H.; Cheuk R.; Hayashizaki Y.; Ishida J.; Jones T.;
Kamiya A.; Karlin-Neumann G.; Kawai J.; Kim C.; Lam B.; Lin J.;
Miranda M.; Narusaka M.; Nguyen M.; Palm C.J.; Sakurai T.; Satou M.;
Seki M.; Shinn P.; Southwick A.; Shinozaki K.; Davis R.W.; Ecker J.R.;
Theologis A.;
"Arabidopsis Full Length cDNA Clones.";
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
Yamada K.; Banh J.; Chan M.M.; Chang C.H.; Chang E.; Dale G.M.;
Deng J.M.; Goldsmith A.D.; Lee J.M.; Onodera C.S.; Quach H.L.;
Tang C.; Toriumi M.; Wu H.C.; Yamamura Y.; Yu G.; Bowser L.;
Carninci P.; Chen H.; Cheuk R.; Hayashizaki Y.; Ishida J.; Jones T.;
Kamiya A.; Karlin-Neumann G.; Kawai J.; Kim C.; Lam B.; Lin J.;
Miranda M.; Narusaka M.; Nguyen M.; Palm C.J.; Sakurai T.; Satou M.;
Seki M.; Shinn P.; Southwick A.; Shinozaki K.; Davis R.W.; Ecker J.R.;
Theologis A.;
"Arabidopsis Open Reading Frame (ORF) Clones.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AC011437; AAF04890.1; -
EMBL; AY080709; AAL85027.1; -
EMBL; AY117275; AAM51350.1; -
GO: 0005524; F:ATP binding; IEA.
GO: 0004830; F:tryptophan-tRNA ligase activity; IEA.
GO: 0006436; P:tryptophanyl-tRNA aminoacylation; IEA.

```



Matches 203; Conservative 66; Mismatches 114; Indels 22; Gaps 5;

QY 9 BAEDFVDPWTVQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQPHHFLRRGIFF 68  
DI 227 ERSKD-VTFWDVNNINBERGINYNKLIKFGGSKITENHIKRIEKLTSKAHFIRRGIF 285  
QY 69 SHRDNOVLDAVENKDPFLYTGSPSEAMHVGLHIFITKMLQDVNPLVITQMTD 128  
DI 286 SHRDDELINLYEHOHKFYITGSPSSLSMHLGLHIFPFCKYLQEAFLVPLVQLSD 345  
QY 129 EKLW-KDLTLDQAYGDAVENAKDIIACGDFINIKTFISDLDYMGSSGFYKXNVKIOKH 187  
DI 346 EKLWQNVSEYINTLNTNENKDIISVGLNPELTFIKNEYAGY---LPTVLSIHKH 402  
QY 188 VTFNQVGIFFGTSDCIGKISFPAIOAAPSNSFPQIFRDRTDIQLIPCAIDQDPYF 247  
DI 403 TTLNQSMNVFGFNHSDNIGKISYFSQIAPCFSCFPNPL--GKNIPCLVPGIDQDPYF 460  
QY 248 RMTDVAIPRIGYPKAPALLHSTFFPALQGAOTKMSASDP-----NSSIFLT 292  
DI 461 RLSRDIIVKALHAPVHVSVFMPCLQSVNMSSTKKKKDNGKSNSTFDDHNSVIFLT 520  
QY 293 DTAQIKTKVNHAFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDKLEQIRKDYTS 352  
DI 521 DTPEQIKNKINKYAFSGGTTIQEHREKGNLKDYSQYLYRLLDCKNLNEIGEKYK 580  
QY 353 GAMLTKELKALIEVLOPLTAHQARKEVTDEIVKSFMTPRKLS 397  
DI 581 GEMLSGEIKLIDVLTELVLHQEKKSSTDEISYFFDPNPKS 625

RESULT 13  
Q8SQY5 PRELIMINARY; PRT; 385 AA.

AC Q8SQY5;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Tryptophanyl tRNA synthetase.  
GN ECUL1.0530.  
OS Eucaryote; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
OX NCBI\_TaxID=6035;  
RN R1  
RP SEQUENCE FROM N.A.  
RA STRAIN=GB-M1;  
RA Genoscope;  
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN R2  
RP SEQUENCE FROM N.A.  
RA STRAIN=GB-M1;  
RA MEDLINE=21576510; PubMed=11719806;  
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat P.,  
RA Frensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,  
RA Delbac P., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
RA Weissenbach J., Vivares C.P.;  
RT "Genome sequence and gene compaction of the eukaryote parasite  
RT Encephalitozoon cuniculi";  
RL Nature 414:450-453(2001).  
RL EMBL; AL590450; CAD25963.1; --  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR PRINTS; PR01039; TRNASYNTHTRP.  
DR TIGRFAMs; TIGR00233; trps; 1.  
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
SQ SEQUENCE 385 AA; 44188 MW; 2865C3AC95FC859 CRC64;

Query Match 44.5%; Score 942; DB 5; Length 385;  
Best Local Similarity 38.1%; Pred. No. 5e-61;

Best Local Similarity 46.1%; Pred. No. 5e-74;  
Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

QY 12 EDFVDPWTVQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQPHHFLRRGIFF 67  
DI 3 EQRITPDVEVWSTDEVPAIDYDKIINQFGCEKFNQALADRLKLSGKPAHVFRRGIV 62  
QY 68 PSHRDMNOVLDAVENKDPFLYTGSPSEAMHVGLHIFITKMLQDVNPLVITQMTD 127  
DI 63 FAHRDFNLLEIANNRPFIYTGSPSSKTMHIGHTIPFLCKYMQDAFKRLV-QITD 122  
QY 128 DEKYLKMDLTLDQAYGDAVENAKDIIACGDFINIKTFISDLDYMGSSGFYKXNVKIOKH 187  
DI 123 DEKFLWKSRLLEDAMAYGRENKIKOIVALGDFPKLTYIESNVE--ASHFENIKIKST 179  
QY 188 VTFNQVGIFFGTSDCIGKISFPAIOAAPSNSFPQIFRDRTDIQLIPCAIDQDPYF 247  
DI 180 ININEAIVKVFQDNSSNIGQVGFPAKEIAPCFSSSFRFIGK---GAMCLVPAAVDQDPFF 236  
QY 248 RMTDVAIPRIGYPKAPALLHSTFFPALQGAOTKMSASDPNSIFLTDTAKQIKTKVNHAF 307  
DI 237 RLARDKAKALGEEKPSSIYVLSLLFDLKGVNRMKMSASDPNSIYLLDCAQPTIRKKTIAAY 296  
QY 308 SGRDRTIEHRQFGNCVDVSMYLTFFLEDDKLEQIRKDYTS GAMLTKELKALIEV 367  
DI 297 SGRKRTLEHREKGGDIDVDVPEYLVXFLDDQDELEKYRSGYIKGETSKEMKEKCVV 356  
QY 368 LOPLIAEHQARKEVTDEIVKSFMTPRK 395  
DI 357 IQEFVSRYSQSRKRVTDLLEAFIDINK 384

RESULT 14  
Q8UIF5 PRELIMINARY; PRT; 430 AA.

AC Q8UIF5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Tryptophanyl tRNA synthetase.  
GN LI063.04.  
OS Leishmania major.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
RN R1  
RP SEQUENCE FROM N.A.  
RA STRAIN=Friedlin;  
RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,  
RA Barrell B.G.;  
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN R2  
RP SEQUENCE FROM N.A.  
RA STRAIN=Friedlin;  
RA MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
RA Smith D.F.;  
RT "A physical map of the Leishmania major Friedlin genome";  
RL Genome Res. 8:135-145(1998).  
RL EMBL; AL121862; CAB58393.1; --  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR PRINTS; PR01039; TRNASYNTHTRP.  
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
KW Aminoacyl-tRNA synthetase.  
SQ SEQUENCE 430 AA; 54904 MW; 172C68622C51D3D CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;  
Best Local Similarity 38.1%; Pred. No. 5e-61;

```

Matches 173; Conservative 87; Mismatches 129; Indels 65; Gaps 11;
C/ 3 HGDATEAEDFDVDPWTVOTSSAKGIDYDKLIVFGSSKID-----KELINRIERAT- 54
D/ 40 HPRDGDAGAEEDVITPVWVAAGPQGINYDRVLITFKAERMDGARGHMRDVMKCRKRTM 99
Q/ 55 -----CORP-----HHFLRGFFFSHRDM 73
D/ 100 TTTPSANAEGVATPVLDVAGDAALQAAASPAMQORQQPPALALHHFFHFDIAFSHRDL 159
Q/ 74 NOVLDAYE-----NKKPFVLYTGRGSPSEAHVGHLPFIPTKWLQDVFNVLVIQMTDD 128
D/ 160 HKALVDIEASVKTGERSVLYTGRGSPSAGTMHLGVLFPMLTKYLQDVFLPLVVIQITDD 219
Q/ 129 EKVLWKDLTLDQAYGDAV--ENAKDIIACGPDINKTIFSDLDYMGSSGFYKNVVKIQK 186
D/ 220 EKFLFDVPFEGAKADELIRSNIKDIIAENFNPRHTFIERNTHYMG---DMYPTVLRQLR 276
Q/ 187 HVTENQVKGIFGFTSDICIGKISFPAIQAPSPNSFPQIFR--DRTOIQCLIPCAIDQD 244
D/ 277 SMGTGNAVKTILGITSNDNVKLAFFATQAPCESTAFRRVLRNGDR-PMRCLIPCAIDQD 335
Q/ 245 PYFRMTRDVAPRIGVPKPALLHSTFFPALQGAQTKMSAS--DPNSSIFLDTAKOIKTKVN 303
D/ 336 PPFVLTAAALRLKQLPPALLHTKPLPALKGLHKMSSSAEKGVTILHDTQVRRKL- 394
Q/ 304 KHAFGSGRDTIEHRQFGGNCVDVNSFMVLTFLFLEDDDKLEQIRKDYTSGLMTGLKKA 363
D/ 395 RRAFGSGCATLEQMOETGANLELDVAYQYLRFFCPDDTLFADVTQYRSGTILNSGEVKDL 454
Q/ 364 LIE-VLQPLIAEHOARRKEVTDIEIVKEFMTPRKL 396
D/ 455 AADCIIRREVLDHWRERRATVTDVDDVVEFCIRDI 488

```

## RESULT 15

```

Q95295
ID Q95295 PRELIMINARY; PRT; 136 AA.
AC Q95295;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trypctophanyl-tRNA synthase (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cdna
  library."
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81267; CAB03585.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
D/ InterPro; IPR002305; tRNA-synt_1b.
D/ InterPro; IPR002306; tRNA-synt_1b.
D/ Pfam; PF00579; tRNA-synt_1b; 1.
D/ PRINTS; PR01039; TRNASYNTHTRP.
F/ NON TER 1
F/ NON TER 136
S/ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

```

Query Match 32.2%; Score 682; DB 6; Length 136;  
 Best Local Similarity 94.1%; Pred No. 7.2e-52;

Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

Q/ 211 PAIQAPSPNSFPQIFRDRDTDIQCLIPCAIDQDPYFRMTRDVAPRIGVYKPKALLHSTFF 270
D/ 1 PAIQAPSPSSFPQIFRDRDTDIQCLIPCAIDQDPYFRMTRDVAPRIGVYKPKALLHSTFF 60

```

```

Q/ 271 PALQGAOTKMSASDPNSSIFLDTAKOIKTKVNHAFSGGSDTIEHRQFGGNCVDVVSF 330
D/ 61 PALQGAOTKMSXXXAPNSSIFLDTAKOIKTKVNHAFSGGSDTIEHRQFGGNCVDVVSF 220
Q/ 331 MYLTFPLEDDDDKLEQI 346
D/ 121 MYLTFPLEDDDDRXEQI 136

```

Search completed: August 24, 2004, 18:03:03  
 Job time : 122 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.  
 OM protein - protein search, using sw model  
 Run on: August 24, 2004, 17:47:41 ; Search time 32 Seconds  
 (without alignments)  
 546.938 Million cell updates/sec  
 Title: US-09-813-718-10\_COPY\_71\_471  
 Perfect score: 2116  
 Sequence: 1 SNHGPD:TEAREDFVDPWTV.....VTDEIVKEFMTPRKLSDFDQ 401  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued Patents AA.\*  
 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
 5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pep.\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result		%		SUMMARIES		Description	
No.	Score	Match	Length	DB	ID		
1	1218.5	57.6	424	3	US-08-876-885-26	Sequence 26, Appl	
2	195	9.2	338	4	US-09-328-352-4201	Sequence 4201, Ap	
3	186	8.8	348	4	US-09-107-532A-5765	Sequence 5765, Ap	
4	185.5	8.8	341	3	US-08-928-100-2	Sequence 2, Appl	
5	185.5	8.8	341	4	US-09-492-581-2	Sequence 2, Appl	
6	185.5	8.8	341	4	US-09-425-666-2	Sequence 2, Appl	
7	174	8.2	335	4	US-09-134-000C-5784	Sequence 5784, Ap	
8	164.5	7.8	409	2	US-08-743-130A-39	Sequence 39, Appl	
9	162.5	7.7	409	2	US-08-743-130A-2	Sequence 2, Appl	
10	157.5	7.4	385	4	US-09-489-039A-8660	Sequence 8660, Ap	
11	152	7.2	344	4	US-09-198-452A-857	Sequence 4, Appl	
12	132.5	6.3	388	2	US-08-705-868-4	Sequence 4, Appl	
13	132.5	6.3	388	3	US-09-123-615-4	Sequence 4, Appl	
14	123	5.8	356	4	US-09-543-681A-4257	Sequence 4257, Ap	
15	118	5.6	349	4	US-09-543-681A-6769	Sequence 6769, Ap	
16	118	5.6	426	4	US-09-107-532A-4046	Sequence 4046, Ap	
17	113.5	5.4	418	3	US-08-855-910-11	Sequence 11, Appl	
18	113.5	5.4	433	4	US-09-134-000C-3686	Sequence 3686, Ap	
19	109	5.2	377	3	US-09-352-990-28	Sequence 28, Appl	
20	108	5.1	197	2	US-08-923-867-2	Sequence 2, Appl	
21	108	5.1	197	3	US-08-928-100-4	Sequence 4, Appl	
22	108	5.1	197	3	US-09-183-134-2	Sequence 2, Appl	
23	108	5.1	197	4	US-09-492-581-4	Sequence 4, Appl	
24	108	5.1	197	4	US-09-425-666-4	Sequence 4, Appl	
25	104.5	4.9	370	2	US-08-415-593-45	Sequence 45, Appl	
26	100.5	4.7	427	4	US-09-134-001C-5141	Sequence 5141, Ap	
27	97.5	4.6	335	4	US-09-489-039A-10919	Sequence 10919, A	

28 95 4.5 418 3 US-08-844-054-2 Sequence 2, Appli  
 29 95 4.5 418 3 US-09-347-333-2 Sequence 2, Appli  
 30 95 4.5 464 4 US-09-134-001C-4701 Sequence 4701, Ap  
 31 94.5 4.5 877 2 US-08-907-166-8 Sequence 8, Appli  
 32 94.5 4.5 877 4 US-09-391-340-8 Sequence 8, Appli  
 33 92.5 4.4 434 4 US-09-543-681A-4577 Sequence 4577, Ap  
 34 91.5 4.3 454 4 US-09-459-039A-9702 Sequence 9702, Ap  
 35 91 4.3 892 4 US-09-585-858-2 Sequence 2, Appli  
 36 91 4.3 892 4 US-09-585-858-2 Sequence 2, Appli  
 37 89.5 4.2 344 3 US-09-393-554-2 Sequence 3, Appli  
 38 89 4.2 1467 4 US-09-134-000C-6740 Sequence 6740, Ap  
 39 88.5 4.2 501 4 US-09-157-257-8 Sequence 8, Appli  
 40 88.5 4.2 898 1 US-08-465-995A-4 Sequence 4, Appli  
 41 88.5 4.2 898 2 US-08-465-994C-4 Sequence 4, Appli  
 42 88.5 4.2 898 2 US-08-966-145-4 Sequence 4, Appli  
 43 88.5 4.2 920 2 US-08-101-593-4 Sequence 3, Appli  
 44 88 4.2 606 2 US-08-883-534-3 Sequence 3, Appli  
 45 88 4.2 606 3 US-09-204-764-3 Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
 US-08-876-885-26  
 ; Sequence 26, Application US/08576885  
 ; Patent No. 6174713

GENERAL INFORMATION:  
 APPLICANT: Shen, Xiaoyu  
 APPLICANT: Eouman, Fariba  
 TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPHANYL-TRNA  
 TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING  
 TITLE OF INVENTION: SAME  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02173

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/876,885  
 FILING DATE: 16-JUN-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook, David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CP197-02  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 861-6240  
 TELEFAX: (781) 861-9540  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 424 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-876-885-26

Query Match 57.6%; Score 1218.5; DB 3; Length 424;  
 Best Local Similarity 57.1%; Pred. No. 4.1e-128;  
 Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 8 TEAREDFVDPWTVQ-----TSSAKGIDYDKLVIFGSGSKIDKELINIERATGQRPHFLR 63

Db 23 TEESEQKITPWEVEGAVDVGKSMGIDYDKLISQGTGKHITETLRFKQVGTGEPHFLK 72



```

QY 64 RGIFFSRDMQVLDAYENKPPVLYTGRGSSSEAMVGHLPFIPTKWLQDVFNWPLVI 123
Db 73 RGVPFSQRLDRLDLYEHGEPEFLYTGSGPSSSMEHGMWPFIFTKLQEVDFWPLVI 132
QY 124 QMTDDKYLWK-DLTDQAYGDAVENAKDIIACGFIINKTFISDLDYMGSSGFYKVV 182
Db 133 ELTDDKFLFKHJLTDVKGFAENAKDIIAGVFNPFIFSDLOQYMG--GAFYENVV 190
QY 183 KIQKHVTFNQVKGIFGFTSDCIGKISFPFAIQAPGFSNSFPQIFRDRDIOCLIPCAID 242
Db 191 RTSRQITTTAKAVGFTSDCIGKIHFAISQIATAPPSSFDVLGLPPTKPLIPCAID 250
QY 243 QDPYFMTVDVAPRIGYKPKALLHSTFFPALQAGATMSASDINSIFLTDIAKQINTKV 302
Db 251 QDPYFVCRDVAQDKRFTKFAHIAKFFPALQAGASTMSASDITTSIFMGDTAKIQKKI 310
QY 303 NKHAFSGGRTTIEHRQFGNCVDVSEMYLTFFLEDDDKLEQIRKDYTSGLMTGELKK 362
Db 311 NKYAFSGGRATAEHRELGNPEVDVAFQLSFFSYDDEKLAQLEQYRKGELSGEMKK 370
QY 363 ALIEVLQPLIAEHQAARKVETDIBIVKFEFTPKLSF 398
Db 371 ECIITVLEFVSAYQERRSKVDQVQVEKPKHLVF 406

```

## RESULT 2

```

US-09-328-352-4201
; Sequence 4201, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4201
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US 09-328-352-4201

```

```

Query Match      9.2%; Score 195; DB 4; Length 338;
Best Local Similarity 25.5%; Pred. No. 4e-13;
Matches 84; Conservative 56; Mismatches 138; Indels 46; Gaps 13;
QY 82 NKDPFYLTYGRGSSSEAMVGHLPFIPTKWLQDVFNWPLVI-----QMTDDKYLWKDL 136
b 6 DQRPILL-TGDRPTGQ-LHLGHFVGSLSRVLGQSHHQHLLLLADAQALTDNAD----- 57
Y 137 TLDQAYGDAVENAKDIIACGFDINKTFI--SDLDYMGSSGFYKVN---VVKIQKHVTFN 191
b 58 NPDKVRNRNILEVALDYLAAGIDPTKTTICVQSCLPALNLTWYLYNFVTVARLERNPTIK 117
Y 192 QVKGIFGFTSDCIGKISFPFAIQAPGFSNSFPQIFRDRDIOCLIPCAIDQDPYFMT 250
b 118 SETQMRGFERDIPAGFLCYPAQAA-----DITAFKATVVPVGVGSDQIPMIQT 165
Y 251 RDVAPRIGYKPKALLHSTFFPALQAGATMSASDINSIFLTDIAKQINTKV 301
b 166 NEIVRRVNRQIGDQLPECKALLSNMARLPQFDG-KAKMSKSLGN-TIVLNASDKOIKKA 223
Y 302 VNHAFSGGRTTIEHRQFGNCVDVSEMYLTFFLEDDDKLEQIRKDYTSGLMTGELKK 361
b 224 VNAMYTDPNHLRIEDPQGVGN-----IVFTYLDADFDPNKEBVEELKAHYRGGGIDGTVK 279
Y 362 KALIEVLQPLIAEHQAARKVETD 385
b 280 KRLEGVLKELITPIRERREBLAKD 303

```

## RESULT 3

```

US-09-107-532A-5765
; Sequence 5765, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5765:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...348
; SEQUENCE DESCRIPTION: SEQ ID NO: 5765:
US-09-107-532A-5765

```

```

Query Match      8.8%; Score 186; DB 4; Length 348;
Best Local Similarity 24.9%; Pred. No. 4.3e-12;
Matches 82; Conservative 57; Mismatches 120; Indels 70; Gaps 14;
QY 88 LYTGGRSSSEAMVGHLPFIPTKWLQDVFNWPLVIQMTDDKYLWKDLTDQAYGDAVE 147
Db 17 ILTGDRPTGK-LHLGHVGSLSKTR-----VAMQADESNQFLVMIADMQALTDNAK 65
QY 148 N-----AKDIIACGFDINKT--FISDLDYMGSSGFYKVV---KIQKHVT-- 189
Db 66 NPEKVVSNVLQVALDYLAAGIDPAKSTLFIQSIQPELAELTWYLYNVSVGRVNRNPTVK 125
QY 190 --FNQVKGIFGFTSDCIGKISFPFAIQAPGFSNSFPQIFRDRDIOCLIPCAIDQDPY 246
Db 126 TEIEQKK-----FGESVPTGFFIYFVSQAA-----DITAFKANLVPGVSDQKPM 169
QY 247 FRMTRD-----VAPRIGYKPKALLHSTFFPALQAGATMSASDINSIFLTD 293
Db 170 LEQTOEIVQSFNHTYGEVLVEPKGVFPKGMGR---LPGIDG-NGRKMSKSLGN-GIYISD 224
QY 234 TAKQIKTKVNKHAFSGGRTTIEHRQFGNCVDVSEMYLTFFLEDDDKLEQIRKDYTSG 353
Db 225 PADVLQKKV-----MSMYTDPNHLHVQDPGQGVGNKXNVTYLDVFTGDKNEA-LEEMKAHYRRG 280

```

QY 354 AMLTGELKALIEVLQPLIAEHQARRKEV 382  
 Db 281 GLGDVKIKRVLIDVLEAFAPIRARREL 303

RESULT 4  
 US-08-928-100-2  
 ; Sequence 2, Application US/08928100  
 ; Patent No. 6046174  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gentry, Danile  
 ; APPLICANT: Greenwood, Claire  
 ; APPLICANT: Lawlor, Elizabeth  
 ; TITLE OF INVENTION: NO. 6046174el trps  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,100  
 ; FILING DATE: 12-SEP-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 9619072.3  
 ; FILING DATE: 12-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gimmi, Edward R.  
 ; REGISTRATION NUMBER: 38,891  
 ; REFERENCE/DOCKET NUMBER: P31624-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-4478  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 341 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-928-100-2

Query Match 8.8%; Score 185.5; DB 3; Length 341;  
 Best Local Similarity 24.0%; Pred. No. 4.8e-12;  
 Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVNVPLVQMTDDKEY-LWKDLTLDOAY 142  
 Db 3 KPIIL-TGDRPTGK-LHGHYVGLSKNR-----VLLQEEKYDMFVLAQQAL 49  
 QY 143 GDAVEN-----AKDIIACGFDINKTIF--SOLDYMGSSGFYKNVY---KIOK 186  
 Db 50 TDHAKDPQTIVESIGNVALDYLAAGLDPNKSTIFQSQIPELAELSMYMNVLARLER 109  
 QY 187 HVTFNQVKGIFGFTSDCIGKISFPAIAQA--PSFSNSFPQIFRDRTOICLIPCAIDQD 244  
 Db 110 NPTVKTEISQKGFSGSIPTGLVYPIAQAADITAFKANY-----VPGTDOX 156  
 QY 245 PYFRMTRD-----VAPRIGPKALLHSTFFPALQGAQTWMSASDPNSSIF 290  
 Db 157 PMIQTRIEIVRSFNAYNCDVLVEPEGIYPENE--RAGRLPLDGG-NAKMSKS-LNNGIY 212  
 QY 291 LTDATKQIKTKVKNHAFSGGRDTIEHROFGNCNDVDVSFMYLTFF--LEDDKLEQIKR 348

Db 213 LAADADTLRKKVMSMYTDPDHRVDEDPGKIEN-----MVFHLDVFGRPEDAQEIADMK 268  
 QY 349 DYTSAMLTGELKALIEVLQPLIAEHQARRKEVTD 386  
 Db 269 RYQGGGLGDKTKRYLLLEILRELGLPIRRIEFAKDM 306

RESULT 5  
 US-09-492-581-2  
 ; Sequence 2, Application US/09492581  
 ; Patent No. 6346409  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gentry, Danile  
 ; APPLICANT: Greenwood, Claire  
 ; APPLICANT: Lawlor, Elizabeth  
 ; TITLE OF INVENTION: NO. 6346409el trps  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/492,581  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,100  
 ; FILING DATE: 12-SEP-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gimmi, Edward R.  
 ; REGISTRATION NUMBER: 38,891  
 ; REFERENCE/DOCKET NUMBER: P31624-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-4478  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 341 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-492-581-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
 Best Local Similarity 24.0%; Pred. No. 4.8e-12;  
 Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVNVPLVQMTDDKEY-LWKDLTLDOAY 142  
 Db 3 KPIIL-TGDRPTGK-LHGHYVGLSKNR-----VLLQEEKYDMFVLAQQAL 49  
 QY 143 GDAVEN-----AKDIIACGFDINKTIF--SOLDYMGSSGFYKNVY---KIOK 186  
 Db 50 TDHAKDPQTIVESIGNVALDYLAAGLDPNKSTIFQSQIPELAELSMYMNVLARLER 109  
 QY 187 HVTFNQVKGIFGFTSDCIGKISFPAIAQA--PSFSNSFPQIFRDRTOICLIPCAIDQD 244  
 Db 110 NPTVKTEISQKGFSGSIPTGLVYPIAQAADITAFKANY-----VPGTDOX 156  
 QY 245 PYFRMTRD-----VAPRIGPKALLHSTFFPALQGAQTWMSASDPNSSIF 290

Db 157 PMIEQREIVRSFNAYNCVLPVEPGIYFENE--RAGRLPGLDG-NAKMSKS-LNNGIY 212  
 QY 291 LDTAKQIKTKVNHAFSGGRDTIEHRQFGNCVDVSMYLTFF--LEDDDKLEQIRK 348  
 Db 213 LADADTLRKVKMSMYTDPDHIRVEDPGKIEGN---MVPHYLDVFGRPEDAQEIADMKE 268  
 QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDEI 386  
 Db 269 RYORGGLGDVKRYLLEILRELGPRIERRIEFAKDM 306

## RESULT 6

US-09-425-666-2

; Sequence 2, Application US/09425666

; Patent No. 6416976

; GENERAL INFORMATION:

; APPLICANT: Gentry, Danile

; APPLICANT: Greenwood, Claire

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: No. 6416976el trps

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/425,666

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/928,100

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Gimmi, Edward R

; REGISTRATION NUMBER: 38,891

; REFERENCE/DOCKET NUMBER: P31624-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-4478

; TELEFAX: 610-270-5090

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 341 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-425-666-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
 Best Local Similarity 24.0%; Pred. No. 4,8e-12;  
 Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPPLYTGRGPSSEAMHVLHPIFTKMLQDVNFVPLVIQMTDDKY-LWKDLTLDQAY 142  
 Db 3 KPILL-TGDRPTGK-LHGHVYVSLNKR-----VILQEEKYDMFVFLAQDQAL 49  
 QY 143 GRAVEN-----AKDIIACGFDINKTIF--SLDLYMGSSGFGYKNVY---KIQK 186  
 Db 50 TDHAKDPQTVESIGNVALDVLAVGLDENKSTFIQSQIPELAEIWSYNNVLSLARLER 109  
 QY 187 HVTFNQVKGIFGTDSCCKGKSFPAIQAA--PSFNSPFIQFRDRTDIQCLIPCAIDQD 244  
 Db 110 NPTVKTISQKFGSGSIPTGFLVPIYIAQAADITAFKANY-----VPGVGDQX 156

QY 245 PYFRMTRD-----VAPRIGYKPKALLHSTFTFPALQGAQTKMSASDPNSSIF 290  
 Db 157 PMIEQREIVRSFNAYNCVLPVEPGIYFENE--RAGRLPGLDG-NAKMSKS-LNNGIY 212  
 QY 291 LDTAKQIKTKVNHAFSGGRDTIEHRQFGNCVDVSMYLTFF--LEDDDKLEQIRK 348  
 Db 213 LADADTLRKVKMSMYTDPDHIRVEDPGKIEGN---MVPHYLDVFGRPEDAQEIADMKE 268  
 QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDEI 386  
 Db 269 RYORGGLGDVKRYLLEILRELGPRIERRIEFAKDM 306

## RESULT 7

US-09-134-000C-5784

; Sequence 5784, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lyne Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; PRIOR FILING DATE: 1998-08-13

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 5784

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

; US-09-134-000C-5784

Query Match 8.2%; Score 174; DB 4; Length 335;  
 Best Local Similarity 23.5%; Pred. No. 9,1e-11;  
 Matches 81; Conservative 53; Mismatches 123; Indels 82; Gaps 16;

QY 88 LYTGEPSSEAMHVGHLIP-----FIFTKMLQDVNFVPLVIQMTDDKY 131  
 Db 5 ILTGDRPTGK-LHGHVYVSLNKRVMQADPTMQLFVMIADLQ-----ALTDNAR- 53  
 QY 132 LWKDLTLDQAYGDAVENAKDIIACGFDINKTIF--SLDLYMGSSGFGYKNVY---KIQK 186  
 Db 54 -----NPEKYSANVLEVALDVLAVGLDPTKTIIFIQSQIPQIAELTMYLNLVITTSVR 108  
 QY 187 HVT-----FNQVKGIFGTDSCICKISFPATQAAPSFNSPFIQFRDRTDIQC-LIPCAI 241  
 Db 109 NPTVKABIEQKK-----FGEQVPTGFFIYVVSQA-----DITAFQANLVVPG 152  
 QY 242 DQDPYFRMTRD-----VAPRIGYKPKALLHSTFTFPALQGAQTKMSASDPNSS 288  
 Db 153 DQKPMLEQAQEI VHSFNQTYGEVLRPEAVLPKGMGR---LPGIDG-NGKMSKSLGN-G 207  
 QY 289 IFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRK 348  
 Db 208 IYLSDPRAEVVQKKVMSMYTDPNHIRIEDPGQVEGN---MVFYLDVFGRPENKEYIELKE 263  
 QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARKEV--TDEIVKRFM 391  
 Db 264 HYRHGGIGLDVKRYLIDVLEELAPIRRERRELAKNPEAIMEML 308

## RESULT 8

US-08-743-130A-39

; Sequence 39, Application US/08743130A

; Patent No. 5871987

; GENERAL INFORMATION:

; APPLICANT: Sassanfar, Mandana

; APPLICANT: Gallant, Paul L.

; APPLICANT: Shen, Xiaoyu

; APPLICANT: Tao, Nianjun

APPLICANT: Tao, Jianshi  
 APPLICANT: Houman, Fariba  
 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/743,130A  
 FILING DATE: 01-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook Esq., David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-12  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 861-6240  
 TELEFAX: (617) 861-9540  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 409 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: linear  
 TOPOLOGY: linear  
 US-08-743-130A-39

Query Match 7.8%; Score 164.5; DB 2; Length 409;  
 Best Local Similarity 20.7%; Pred. No. 1.5e-09;  
 Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;  
 QY 76 VLDAYENK-KPFVLYTGRGSSSEAMHVGHLPFI-----FTKWLDV-----F 117  
 Db 27 IKDVLKERNRPVKIYWGTAFTGPK-HCGYFVPMKLAHFLKAGEVTVLLADLHAFLDNM 85  
 QY 118 NVPLVIQMTDDEKYLKDLTDOAYGDAVENAKDIACGDFINKTFESLDLYMGSSGF 177  
 Db 86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGDYV---MDLFKLSNIV 142  
 QY 178 YKNVVK-----IQKHVTFNQVKGIFGFTDSDCIGKISFPALQAPSFNSFPQIFDRD 232  
 Db 143 SQNDKRAKAGADVVKQVANPLLSGLI-----YPLMQA-----IDEEHLG 180  
 QY 233 IQCLIPCAIDQDPYFMTDVAIRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLT 292  
 Db 181 VDAQFG-GVDDQKIFVLAEENLPSIGYKRAHLMNPMVPG-L-GQGGKMSASDPNSKIDII 238  
 QY 293 DTAKOIKTKVKNHAFSGG--RDT-----IEHRQFGGN 323  
 Db 239 EEPKVVKKKNSAYCAPGELKDNGLIAFYVIQPIAEKLTGVEGAFKLDIDRPEKYGG- 297  
 QY 324 CDVDSFMYLTFLEDDDDKLEQIRKDYTSGLMTGLKXALI-----EVLQPLIAEHQAR 379  
 Db 298 ---PLSY-----DSIEQLKADFVUGKLAAPDLKGSVADKINELLAPIRAEFESS- 343  
 QY 382 KEVTEIVKEMTPRK 395  
 Db 344 -----EEFQVQAK 351

RESULT 9

US-08-743-130A-2

Sequence 2, Application US/08743130A

Patent No. 5871987

GENERAL INFORMATION:  
 APPLICANT: Saseanfar, Mandana  
 APPLICANT: Gallant, Paul L.  
 APPLICANT: Shen, Xiaoyu  
 APPLICANT: Tao, Nianjun  
 APPLICANT: Tao, Jianshi  
 APPLICANT: Houman, Fariba  
 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/743,130A  
 FILING DATE: 01-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook Esq., David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-12  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 861-6240  
 TELEFAX: (617) 861-9540  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 409 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: linear  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-743-130A-2

Query Match 7.7%; Score 162.5; DB 2; Length 409;  
 Best Local Similarity 20.7%; Pred. No. 2.5e-09;  
 Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;  
 QY 76 VLDAYENK-KPFVLYTGRGSSSEAMHVGHLPFI-----FTKWLDV-----F 117  
 Db 27 IKDVLKERNRPVKIYWGTAFTGPK-HCGYFVPMKLAHFLKAGEVTVLLADLHAFLDNM 85  
 QY 118 NVPLVIQMTDDEKYLKDLTDOAYGDAVENAKDIACGDFINKTFESLDLYMGSSGF 177  
 Db 86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGDYV---MDLFKLSNIV 142  
 QY 178 YKNVVK-----IQKHVTFNQVKGIFGFTDSDCIGKISFPALQAPSFNSFPQIFDRD 232  
 Db 143 SQNDKRAKAGADVVKQVANPLLSGLI-----YPLMQA-----IDEEHLG 180  
 QY 233 IQCLIPCAIDQDPYFMTDVAIRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLT 292  
 Db 181 VDAQFG-GVDDQKIFVLAEENLPSIGYKRAHLMNPMVPG-L-GQGGKMSASDPNSKIDII 238  
 QY 293 DTAKOIKTKVKNHAFSGG--RDT-----IEHRQFGGN 323  
 Db 239 EEPKVVKKKNSAYCAPGELKDNGLIAFYVIQPIAEKLTGVEGAFKLDIDRPEKYGG- 297  
 QY 324 CDVDSFMYLTFLEDDDDKLEQIRKDYTSGLMTGLKXALI-----EVLQPLIAEHQAR 379  
 Db 298 ---PLSY-----DSIEQLKADFVUGKLAAPDLKGSVADKINELLAPIRAEFESS- 343  
 QY 380 KEVTEIVKEMTPRK 395  
 Db 344 -----EEFQVQAK 351

```

RESULT 10
US-09-489-039A-8660
; Sequence 8660, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.3004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 1999-01-29
; SEQ ID NO 8660
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8660

Query Match      7.4%; Score 157.5; DB 4; Length 385;
Best Local Similarity 24.4%; Pred. No. 8.3e-09;
Matches 87; Conservative 53; Mismatches 142; Indels 75; Gaps 17;

QY 60 HFLRR---GIFFSHRDMQVLDAYENKX-----PFYLYTGRGSPSEAHVGHLPPIFT 110
DB 25 HFQSRKAAVFSGR-----LSAQSDKESLMNHPTILTGRPTGQ-LHLGHVYVGLRQ 78
QY 111 KWLQDVNVPLVIQMTDDEKYLWKDLTLQDAYGDAVENAK-----DIACGFDI 159
DB 79 R-----VALQHNHQFILIADL---QGLTNGSNPKISHHILEVMADYLAUGLDT 126
QY 160 NKTFIF--SDLYNMGSSGFYKNVY---KIQKHVTFNOVKIGIFGTFSDCIGKISFPAIQ 214
DB 127 RLTTICQSLPALAEALSALYNNIVTVARVERNPTKNEIAQKGFARSLPVGFLAYPISQ 185
QY 215 AAPFSFNSFPQIFRDRDTIQCLIPCALDQDPYFRTMDVA-----PRIGYPPKAL 264
DB 187 AA-----DITAFKALVPVGDQDLMIEQNEIVHKNSLTGEPVLSHCKALL 234
QY 265 LHSTFPALQAGTQKMSADPNSIFLDTAKQIKTKVKNHAFSGGRDTIEHRQFG--G 322
DB 235 SEVSRLEPGVDG-NAKMSKSLGN-TLTSATEEEI-----HNAVY-AMYTPTHLRVSDPG 286
QY 323 NCDVDVSEMYLTFFLEDDDKLEQIRKDYTSGLMTGELKALIEVLQPLIAEHQARR 379
DB 287 HVEGNVVFYLDFAHSDKARVAEMKTHYQRGGLGDRQCKNELETCLQTLLAPIRER 343

RESULT 11
US-09-198-452A-857
; Sequence 857, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 857
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-857

Query Match      7.2%; Score 152; DB 4; Length 344;
Best Local Similarity 23.1%; Pred. No. 2.8e-08;
Matches 79; Conservative 50; Mismatches 127; Indels 86; Gaps 16;

```

```

QY 82 NKKPFVLYTGRGSPSEAHVGHLPPIFTFKWLQDVNVPLVIQMTDEKY----- 131
DB 2 NKKKVJ--TGDRTGK-LHLGH-----WVGSIRN-----RLELQNSPEYDCFFIADLH 48
QY 132 -----LWKOLTLQDAYDAVEN-----AKDIIACGFDINKFTIFSDDLQYMGSSGFYKNV 181
DB 49 TLTKIRKEVLQ-----VDNHIYEVLADWLSVGIDFTKSIY-----LQSAPEIYELH 98
QY 182 VKIQKHVTFNOVKIGIFGTFD-----SDCIGKISFPAIQAPFSFNSFPQIFRDRT 231
DB 99 LIPSMJISINRYMGIPSLKDMARNASIEGSLSYGLIGYPILOSA----- 143
QY 232 DIQC-----LIPCALDQDPYFRTMDVA---YKPAALLHS-----TFPPALQAGTQK 279
DB 144 DILLAKAQFVPVGNKNEAHVELTRD--ARNFNLYQGVFPEPEVLOQGEJ--SLVGIDG--QKG 202
QY 280 MSADPNSIFLDTAKQIKTKVKNHAFSGGRDTIEHRQFGGNCDDVDVSEMYLTFFLEDD 339
DB 203 MSKS-ANNAIYLSDSDATITEKVRKMYTDPENRIRATTPGVEGN-----PLFIYHDIENH 257
QY 340 DDKLEQIRKDYTSGLMTGELKALIEVLQPLIAEHQARRKE 381
DB 258 KDEVEEFKARYQGCIOKIDEVKARLAELIHLFKPIKERRSE 299

RESULT 12
US-08-705-868-4
; Sequence 4, Application US/08705868
; Patent No. 5885798
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Synn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,868
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1184699
US-08-705-868-4

Query Match      6.3%; Score 132.5; DB 2; Length 388;
Best Local Similarity 20.8%; Pred. No. 5.4e-06;

```





Db 129 ERNPTVKHELLOK-----NLSRSLPAGELTPVVSQAADITAFSAD-----IV 170  
Qy 238 PCAIDQDPYFRMTRDVAPR-----IGYP-----KPALLHSTFPFALQGAQTWMSADPNSS 288  
Db 171 PAGEDQLEPMIBOTNEIVTKINSLIQCPVLTCKVVGQVGRPLPGTDS-GRMSKSLGN-T 228  
Qy 289 IFLTDTAQIKITKNKHPGSGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEQIRK 348  
Db 229 INLSSTADEIK-----KAVYSMTDPQHIDVASPGHIEGNVFTYLDACQDKATVTAMKA 284  
Qy 349 DYTSGAMLTGELK-KALI-----EVLQPL-----IAEQARRKEVTDE 385  
Db 285 HYQGGCL--GDMCKAMLDILOELLQPIREKRAQLINDKAYLLQVIXEGSDKAKEVTQQ 342  
Qy 386 IVKE 389  
Db 343 KLDE 346

RESULT 15  
US-09-543-681A-6769  
; Sequence 6769, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543.681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6769  
; LENGTH: 349  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6769

Query Match 5.6%; Score 118; DB 4; Length 349;  
Best Local Similarity 21.0%; Pred. No. 0.0002;  
Matches 80; Conservative 46; Mismatches 107; Indels 148; Gaps 20;  
Qy 83 KK3FVLYTGRGSSSEAMHVGHLPFTKWLQDVFNVLVIQMTDDEKYLW-----KDLT 137  
Dt 17 QKPI-VFSGAQESGE-LTIGNYMGAL-ROW-----VQMDDYDCYCIWDQHAIT 63  
Qy 138 LDQAYGDVAVENAKD-----IIACGFDINKTFIFSDLDYMGSSGFKYKVVYKIQKHYT----- 189  
Dt 64 VQDPTFLKRTLDLTALYIACIDPEKSTIF-----VQSHVPQHAQ 105  
Qy 190 -----FNQVKGIFGFTSDC-----IGKISFPFQIAAPSFSNFPQIFRDRDT 232  
Dt 106 LGWALNCVTFYFGLSRMTQFKDKSARHAENINAGLPDYPVLMAA-----DILLYQTN 157  
Qy 233 LQCLIPCALDQDPYFRMTRDVAPRIG--YKPAALLHSTFPF-----ALQGAQTWMSAS 283  
Dt 158 ---QPVVGIDQKHLELSRDIAQRFNAIYGDIFTVPDPPIPKGARVMALQDPFAKXSKS 214  
Qy 284 DPNSS---IFLTD---TAKQIKTKV-----NKEAFSG-----GRDTI 314  
Db 215 DNNRNVIALLDPKAAKKIKRAVTDSEPPRVAYDLENKAGVSNMLDLILAGVTGKTIP 274  
Qy 315 BEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAE 374  
Db 275 ELEAEFECK-----MY-----GHLKGAVAEAVSDMLTN 302  
Qy 375 HOAR-----RKEVTDEIVKE 389  
Db 303 IQERFK--PRNDEALINKIMKE 323

Search completed: August 24, 2004, 18:04:24

Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:18:04 ; Search time 126 Seconds  
(without alignments)  
899.219 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471  
Perfect score: 2116  
Sequence: 1 SNHGPDATBAEDFVDPWTV.....VTDEIVKFEFTPRKLSDFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Pcst-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A: Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	401	5	AAG79549 TrpRS T1
2	2116	100.0	415	4	AAB47617 Human sup
3	2116	100.0	415	5	AAB47617 Human sup
4	2116	100.0	415	5	AAG79548 His6-tag
5	2116	100.0	415	6	ABU72386 Human sup
6	2116	100.0	437	4	AAB47616 Human min
7	2116	100.0	437	5	AAB47616 Human min
8	2116	100.0	437	5	AAG79547 His6-tag
9	2116	100.0	437	6	ABU72385 Human min
10	2116	100.0	484	4	AAB47615 Human ful
11	2116	100.0	484	5	AAB47615 Human ful
12	2116	100.0	484	5	AAG79546 Full leng
13	2116	100.0	484	6	ABU72384 Human ful
14	2101	99.3	471	7	ABU72384 Human pro
15	2101	99.3	471	8	ABU72384 Human pro
16	2101	99.3	471	8	ABU72384 Human pro
17	2096	99.1	471	2	AAU05372 Human HCM
18	1388	94.0	378	5	AAG79541 TrpRS T2
19	1388	94.0	392	4	AAB47618 Human ita
20	1388	94.0	392	5	AAB47618 Human ita
21	1388	94.0	392	5	AAG79544 His6-tag
22	1388	94.0	392	6	ABU72387 Human min
23	1373	93.2	378	6	ABU72387 Human min
24	1338	91.6	475	7	ABU79825 Mouse put
25	1370.5	64.8	430	4	ABB67203 Drosophil

26	1370.5	64.8	430	4	ABB67203	Abb67203 Drosophil
27	1304	61.6	402	3	AAG23698	Aag23698 Arabidops
28	1304	61.6	426	3	AAG23697	Aag23697 Arabidops
29	1218.5	57.6	424	4	AAB66931	Aab66931 Tryptopha
30	1200.5	56.7	424	5	ABP73795	Abp73795 Candida a
31	1163	55.0	432	6	ABR53898	AbR53898 Protein s
32	1125	53.2	433	6	ABJ26487	Abj26487 Aspergill
33	1107.5	52.3	456	7	ABU70160	Abu70160 C. neofo
34	968	45.7	292	3	AAG23699	Aag23699 Arabidops
35	831	39.3	173	3	AAB58517	Aab58517 Lung can
36	803	37.9	385	4	AAB96409	Aab96409 Putative
37	414.5	19.6	179	6	ABJ25887	Abj25887 Aspergill
38	328	15.5	85	5	AAE13515	Aae13515 Human try
39	328	15.5	85	6	ABU72416	Abu72416 Human try
40	292	13.8	85	6	ABU72417	Abu72417 Bovine tr
41	273.5	12.9	85	6	ABU72419	Abu72419 Rabbit tr
42	263	12.4	85	6	ABU72413	Abu72413 Mouse try
43	226.5	10.7	142	3	AAB58219	Aab58219 Lung can
44	203.5	9.6	341	5	ABP26964	Abp26964 Streptoco
45	195	9.2	337	6	ABU17282	Abu17282 Protein e

ALIGNMENTS

RESULT 1  
AAG79549  
ID AAG79549 standard; protein; 401 AA.

XX AAG79549;  
AC AAG79549;  
DT 10-DEC-2002 (first entry)  
XX TrpRS T1 polypeptide.  
DE DE  
XX T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation;  
KW neovascular eye disease; diabetes; age-related macular degeneration;  
KW ocular complication; diabetes; rubecotic glaucoma; retinopathy;  
KW prematurity; keratitis; ischaemic retinopathy; sickle cell;  
KW pathological myopic; ocular histoplasmosis; pterygia; TI;  
KW punctate innerchoriopathy; retinal degeneration; growth factor;  
KW vascularisation; vascular endothelial cell function; angiogenesis.

XX Homo sapiens.

XX WO200267970-A1.

XX 06-SEP-2002.

XX 22-FEB-2002; 2002WO-US005185.

XX 23-FEB-2001; 2001US-0270951P.

XX (SRI ) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K, Friedlander M;

XX WPI; 2002-698635/75.

PT New polypeptides derived from human tryptophanyl-tRNA synthase, useful  
for inhibiting ocular neovascularization in a patient, or for treating  
neovascular eye diseases, e.g. rubecotic glaucoma, retinopathy, keratitis,  
or pterygia.

PS Example 1; Page 78-79; 83pp; English.

XX This sequence represents a novel cleavage product, T1, of recombinant  
human tryptophanyl-tRNA synthase (TrpRS). A related cleavage product, T2,  
is water soluble and comprises residues 94-471 of full length TrpRS. The  
water-soluble T2 polypeptide is useful for inhibiting ocular  
neovascularisation in a patient. The T2 polypeptide is useful for  
treating neovascular eye diseases, e.g. age-related macular degeneration,  
ocular complications of diabetes, rubecotic glaucoma, retinopathy of

CC prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell),  
 CC pathological myopia, ocular histoplasmosis, pterygia, or punitate  
 CC innerchoroidopathy. This polypeptide is particularly useful for treating  
 CC retinal degeneration to prevent the damaging effects of trophic and  
 CC growth factors, and for promoting vascularisation to retard retinal  
 CC degeneration by enhancing blood flow to cells. These are also useful for  
 CC regulating vascular endothelial cell function, and in particular, for  
 CC inhibiting angiogenesis  
 XX  
 XX Sequence 401 AA;

Query Match 100.0%; Score 2116; DB 5; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 Db 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 QY 61 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 120  
 Db 61 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 120  
 QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 Db 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 QY 181 VKIQRHVTNFQVKGIFGFTSDCIKISFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 240  
 Db 181 VKIQRHVTNFQVKGIFGFTSDCIKISFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 240  
 QY 241 IDQDPYFMTROVAPRIGYKPKALHSTFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 300  
 Db 241 IDQDPYFMTROVAPRIGYKPKALHSTFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 300  
 QY 301 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 Db 301 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401  
 Db 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401

RESULT 2  
 AAB47617  
 D AAB47617 standard; protein; 415 AA.  
 X AAB47617;

07-JAN-2002 (first entry)  
 Human supermini TrpRS.

Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain;  
 vascular endothelial cell function; burn; plastic surgery; abdomen;  
 polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 dermal ulcer; diabetic ulcer; endothelialization;  
 cryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

Homo sapiens.  
 WC200174841-A1.  
 11-OCT-2001.

21-MAR-2001; 2001WO-US008966.  
 31-MAR-2000; 2000US-0193471P.

(Scripps Res Inst.

PI

Schimmel P, Wakasugi K;

XX

WPI; 2001-626377/72.

DR

N-PSDB; AAB47617.

XX

New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

PT

vascular endothelial function, in particular for regulating angiogenesis,

PT

tumor metastasis and treating myocardial infarction.

XX

Disclosure; Page 129-30; 150pp; English.

XX

The sequences given in AAB47615-18 show full length and truncated

CC

versions of tyrosyl-tRNA synthetase (TrpRS). The truncated TrpRS of

CC

the invention comprises a Rossmann fold nucleotide binding domain, and is

CC

capable of regulating vascular endothelial cell function. It is of

CC

approx. 40 kDa Dalton molecular weight and is produced by cleavage of

CC

full length TrpRS with polymorphonuclear leucocyte elastase. Truncated

CC

TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing

CC

angiogenesis to a graft, treating myocardial infarction, solid tumor, and

CC

a condition that would benefit from increased or decreased angiogenesis

CC

in a mammal in particular humans. It is also useful in diagnosis and as

CC

a wound healing agent for treating wounds such as dermal ulcers, diabetic

CC

ulcers, burns and injuries and in plastic surgery when reconstruction is

CC

required following a burn or for cosmetic purposes. It is particularly

CC

useful in the treatment of abdominal wounds where there is high risk of

CC

infection. Truncated TrpRS promotes endothelialization in vascular graft

CC

surgery and is used in conjunction with angiography to administer the

CC

angiogenic trRNA synthetase polypeptides or polynucleotides directly to

CC

the lumen and wall of the blood vessel

XX

Sequence 415 AA;

QY

Query Match 100.0%; Score 2116; DB 4; Length 415;

Best Local Similarity 100.0%; Pred. No. 1.1e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60

Db

2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 61

QY

61 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 120

Db

62 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 121

QY

121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180

Db

122 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181

QY

181 VKIQRHVTNFQVKGIFGFTSDCIKISFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 240

Db

182 VKIQRHVTNFQVKGIFGFTSDCIKISFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 241

QY

241 IDQDPYFMTROVAPRIGYKPKALHSTFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 300

Db

242 IDQDPYFMTROVAPRIGYKPKALHSTFPALQAAQPSFNSFPQIFRDTDIQCLIPCA 301

QY

301 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360

Db

302 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 361

QY

361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401

Db

362 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 402

RESULT 3

AAB47617

ID AAB47617 standard; protein; 415 AA.

XX

AC AAB47617;

DT 12-FEB-2002 (first entry)

XX

DE Human supermini tryptophanyl t-RNA synthetase in pBT20B.

XX Human: tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;  
KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;  
KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
KW angiography; gene therapy; tumour; inflammation; vascular permeability;  
KW rheumatoid arthritis; psoriasis; diabetes; diabetic retinopathy.

IX Homo sapiens.

CS WC200175078-A1.

IN 11-OCT-2001.

IF 21-MAR-2001; 2001WO-US008975.

JX 31-MAR-2000; 2000US-0193471P.

IR (SCRI ) SCRIPPS RES INST.

IA Schimmel P, Wakasugi K;

IF WPI; 2002-010784/01.

IR N-PSDB; A022484.

XX Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
FT regulating vascular endothelial cell function, preferably angiogenesis,  
FT is useful for treating solid tumor or suppressing tumor metastasis in  
FT mammal.

PS Example 1; Page 129-130; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly  
CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
CC fold nucleotide binding domain and polynucleotides encoding them. The  
CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
CC sequences are useful for regulating vascular endothelial cell function,  
CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
CC healing agents for re-vascularising damaged tissues. They are useful for  
CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
CC sequences can also be used in plastic surgery when reconstruction is  
CC required following a burn, other trauma, or even for cosmetic purposes.  
CC Angiogenic TrpRS is also used in association with surgery and following  
CC the repair of cuts, for promoting endothelialisation in vascular graft  
CC surgery and for repairing the damage of myocardial infarction and in  
CC conjunction with coronary bypass surgery by stimulating the growth of  
CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
CC immunoassays to detect the presence of tumours. They are also useful for  
CC blocking endogenous angiogenic activity and retard the growth of solid  
CC tumours. These antibodies may also be used to treat inflammation caused  
CC by increased vascular permeability. Inhibiting the activity of TrpRS by  
CC antisense technology is useful for preventing further growth or even  
CC regress solid tumours, and for treating rheumatoid arthritis, psoriasis,  
CC diabetic retinopathy, all of which are characterised by abnormal  
CC angiogenesis. The present sequence is human truncated tryptophanyl t-RNA  
CC synthetase (supermini TrpRS; residues 71-471 of full-length TrpRS  
CC protein) in pBT20B

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 5; Length 415;  
Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGCRPH 60  
DB 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGCRPH 60  
CY 61 FLRRGIFPSHRDMMQVLDAYENKPPFYLYTGRPSSEAMHVGHLLIPFTFKWLQDVFNVP 120  
FLRRGIFPSHRDMMQVLDAYENKPPFYLYTGRPSSEAMHVGHLLIPFTFKWLQDVFNVP 120

DB 62 FLRRGIFPSHRDMMQVLDAYENKPPFYLYTGRPSSEAMHVGHLLIPFTFKWLQDVFNVP 121  
QY 121 LVIQMTDDEKYLWLDLTLDQAYGDAVENAKELIACGFDINKTFFISDLDYMGSSGFYKN 180  
DB 122 LVIQMTDDEKYLWLDLTLDQAYGDAVENAKELIACGFDINKTFFISDLDYMGSSGFYKN 181  
QY 181 VVKIQKHVTNFQVKGIFGFTSDDCIGKISFPAIQAAPSFNSFPOIFRDRDTDIQCLIPCA 240  
DB 182 VVKIQKHVTNFQVKGIFGFTSDDCIGKISFPAIQAAPSFNSFPOIFRDRDTDIQCLIPCA 241  
QY 241 IDQDFYRMTDRVAPRIGYKPKALHSTFFPALQCAQTMSASDPNSSIFLTDTAQIKT 300  
DB 242 IDQDFYRMTDRVAPRIGYKPKALHSTFFPALQCAQTMSASDPNSSIFLTDTAQIKT 301  
QY 301 KVNKHAESGGRTIEHRQFGNCNDVSVFMYLFFLEDDDKLQIRKDYTSGAMLTGEL 360  
DB 302 KVNKHAESGGRTIEHRQFGNCNDVSVFMYLFFLEDDDKLQIRKDYTSGAMLTGEL 361  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
DB 362 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 402

RESULT 4

AAG79548

ID AAG79548 standard; protein; 415 AA.

AC AAG79548;

XX 10-DEC-2002 (first entry)

XX His6-tagged TrpRS T1 polypeptide.

XX T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation;  
KW neovascular eye disease; age-related macular degeneration;  
KW ocular complication; diabetes; rubecotic glaucoma; retinopathy;  
KW prematurity; keratitis; ischaemic retinopathy; sickle cell;  
KW pathological myopia; ocular histoplasmosis; pterygia; T2;  
KW puminate innerchoroidopathy; retinal degeneration; growth factor;  
KW vascularisation; vascular endothelial cell function; angiogenesis.

XX Homo sapiens.

XX WO200267970-A1.

XX 06-SEP-2002.

XX 22-FEB-2002; 2002WO-US005185.

XX 23-FEB-2001; 2001US-0270951P.

XX (SCRI ) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K, Friedlander M;

XX WPI; 2002-698635/75.

XX N-PSDB; ABA00330.

XX New polypeptides derived from human tryptophanyl-tRNA synthase, useful  
FT for inhibiting ocular neovascularization in a patient, or for treating  
FT neovascular eye diseases, e.g. rubecotic glaucoma, retinopathy, keratitis,  
FT or pterygia.

XX Example 1; Page 71-72; 83pp; English.

XX This sequence represents a His6-tagged cleavage product, T1, of  
CC recombinant human tryptophanyl-tRNA synthase (TrpRS). A related cleavage  
CC product, T2, is water soluble and comprises residues 94-471 of full  
CC length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting  
CC ocular neovascularisation in a patient. The T2 polypeptide is useful for  
CC treating neovascular eye diseases, e.g. age-related macular degeneration,  
CC ocular complications of diabetes, e.g. age-related glaucoma, retinopathy of  
CC prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell),

C: pathological myopic, ocular histoplasmosis, pterygia, or punitate  
 C: innerchoroidopathy. This polypeptide is particularly useful for treating  
 C: retinal degeneration to prevent the damaging effects of trophic and  
 C: growth factors, and for promoting vascularisation to retard retinal  
 C: degeneration by enhancing blood flow to cells. These are also useful for  
 C: regulating vascular endothelial cell function, and in particular, for  
 C: inhibiting angiogenesis  
 X: Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 5; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 D: 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 61  
 Q: 61 FLRRGIFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 120  
 D: 62 FLRRGIFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 121  
 Q: 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 D: 122 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 Q: 181 VKIQKHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240  
 D: 182 VKIQKHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 241  
 Q: 241 IDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 300  
 D: 242 IDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 301  
 Q: 301 KVNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 D: 302 KVNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
 Q: 361 KXALIEVLOPLIAEHOARRKEVTDEIVKEFTPRKLSGDFQ 401  
 D: 362 KXALIEVLOPLIAEHOARRKEVTDEIVKEFTPRKLSGDFQ 402

RESULT 5

ABU72386

ID ABU72386 standard; protein; 415 AA.

XX AC ABU72386;

XX DT 16-JUN-2003 (first entry)

XX XX Human supermini Tryptophanyl tRNA synthetase/His tag.

DE DE Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytostatic;

XX KW vulnery; Rossmann fold nucleotide binding domain; chemokine; EMAP II;

XX KW vascular endothelial cell; solid tumour; myocardial infarction; enzyme;

XX KW endothelial monocyte-activating polypeptide II; tumour metastasis;

XX KW wound healing; dermal ulcer; endothelialisation; vascular graft surgery;

XX KW abdominal wound; coronary bypass surgery; gene therapy.

XX OS Homo sapiens.

OS Synthetic.

XX US2002182666-A1.

XX 35-DEC-2002.

XX 21-MAR-2001; 2001US-00813718.

XX 21-MAR-2001; 2001US-00813718.

XX (SCHI/) SCHIMMEL P.

XX (WAKA/) WAKASUGI K.

XX

PI

XX

DR

DR

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Schimmel P, Wakasugi K;

WPI; 2003-340974/32.

N-PSDB; ACA64107.

New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 Rossmann fold nucleotide binding domain or having chemokine activity  
 useful for e.g. for regulating angiogenesis and for treating myocardial  
 infarction.

Example 1; Page 50-51; 91EP; English.

The invention relates to an isolated polypeptide comprising a truncated  
 tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rossmann  
 fold nucleotide binding domain or having chemokine activity. The isolated  
 polypeptide is capable of regulating vascular endothelial cell function.  
 TrpRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 -activating polypeptide II, a proinflammatory cytokine)-like domain and  
 is similar in sequence to TrpRS. Also included are a polynucleotide  
 encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 5' and 3' deletions of the TrpRS polynucleotide, a recombinant vector  
 comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 TrpRS antibody, producing truncations of TrpRS by treating it with a  
 protease, and suppressing angiogenesis, solid tumours or a condition that  
 would benefit from decreased angiogenesis in a mammal by administering to  
 the mammal the composition comprising the TrpRS truncated protein. The  
 isolated polypeptide is useful for the preparation of a pharmaceutical  
 composition for transdermal, transmucosal, enteral or parenteral  
 administration. The truncated tRNA synthetase polypeptide is useful for  
 research, diagnostic, prognostic and therapeutic applications. The tRNA  
 synthetase are useful for regulating vascular endothelial cell function,  
 particularly for regulating angiogenesis, for treating myocardial  
 infarction and solid tumour, and for suppressing tumour metastasis.  
 Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 agents or for treating full thickness wounds such as dermal ulcers, in  
 the promotion of endothelialisation in vascular graft surgery, in the  
 treatment of abdominal wounds where there is a high risk infection, in  
 conjunction with coronary bypass surgery by stimulating the growth of the  
 transplanted tissue, and in gene therapy. The present sequence is a Trp  
 tRNA synthetase protein (full length, truncated or mutant) with a His  
 affinity tag

Sequence 415 AA;

Query Match

Best Local Similarity 100.0%; Score 2116; DB 5; Length 415;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60

D 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 61

QY 61 FLRRGIFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 120

D 62 FLRRGIFFSHRDMNQVLDAVENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 121

QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180

D 122 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181

QY 181 VKIQKHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240

D 182 VKIQKHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 241

QY 241 IDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 300

D 242 IDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 301

QY 301 KVNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360

D 302 KVNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361

QY	361	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	401
Db	61	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	120
Db	84	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	143
QY	121	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	180
Db	144	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	203
QY	181	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	240
Db	204	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	263
QY	241	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	300
Db	264	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	323
QY	301	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	360
Db	324	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	383
QY	361	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	401
Db	384	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	424
RESULT 7			
AAE13492			
ID	AAE13492	standard; protein; 437 AA.	
XX	AAE13492;		
XX	12-FEB-2002	(first entry)	
DE	Human mini tryptophanyl t-RNA synthetase in pET20B.		
XX	Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;		
KW	TyrRS; vascular endothelial cell function; angiogenesis; wound healing;		
KW	re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;		
KW	diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;		
KW	angiography; gene therapy; tumour; inflammation; vascular permeability;		
KW	rheumatoid arthritis; psoriasis; diabetic retinopathy.		
XX	Homo sapiens.		
XX	WO200175078-A1.		
PN	11-OCT-2001.		
XX	21-MAR-2001; 2001WC-US008975.		
XX	31-MAR-2000; 2000US-0193471P.		
XX	(SRI ) SCRIPPS RES INST.		
XX	Schimmel P, Wakasugi K;		
XX	WPI; 2002-010784/01.		
XX	N-PSDB; AAB22483.		
PT	Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of		
PT	regulating vascular endothelial cell function, preferably angiogenesis,		
PT	is useful for treating solid tumor or suppressing tumor metastasis in		
PT	mammal.		
XX	Example 1; Page 123-124; 149pp; English.		
XX	The patent discloses human aminocyl tRNA synthetases, particularly		
CC	truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann		
CC	fold nucleotide binding domain and polynucleotides encoding them. The		
CC	invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS		
CC	sequences are useful for regulating vascular endothelial cell function,		

QY	361	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	401
Db	61	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	120
Db	84	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	143
QY	121	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	180
Db	144	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	203
QY	181	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	240
Db	204	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	263
QY	241	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	300
Db	264	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	323
QY	301	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	360
Db	324	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	383
QY	361	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	401
Db	384	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	424
RESULT 6			
AAE13492			
ID	AAE13492	standard; protein; 437 AA.	
XX	AAE13492;		
XX	11-OCT-2001.		
XX	21-MAR-2001; 2001WC-US008966.		
XX	31-MAR-2000; 2000US-0193471P.		
XX	(SRI ) SCRIPPS RES INST.		
XX	Schimmel P, Wakasugi K;		
XX	WPI; 2001-626377/72.		
XX	N-PSDB; AAB43603.		
PT	New human truncated tyrosyl-tRNA synthetase polypeptide for regulating		
PT	vascular endothelial function, in particular for regulating angiogenesis,		
PT	tumor metastasis and treating myocardial infarction.		
PS	Disclosure; Page 123-24; 150pp; English.		
XX	The sequences given in AAB47615-18 show full length and truncated		
CC	versions of trptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of		
CC	the invention comprises a Rossmann fold nucleotide binding domain, and is		
CC	capable of regulating vascular endothelial cell function. It is of		
CC	approx. 40 kilo Dalton molecular weight and is produced by cleavage of		
CC	full length TrpRS with polymorphonuclear leucocyte elastase. Truncated		
CC	TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing		
CC	angiogenesis to a graft, treating myocardial infarction, solid tumor, and		
CC	a condition that would benefit from increased or decreased angiogenesis		
CC	in a mammal, in particular humans. It is also useful in diagnosis and as		
CC	a wound healing agent for treating wounds such as dermal ulcers, diabetic		
CC	ulcers, burns and injuries and in plastic surgery when reconstruction is		
CC	required following a burn or for cosmetic purposes. It is particularly		
CC	useful in the treatment of abdominal wounds where there is high risk of		
CC	infection. Truncated TrpRS promotes endothelialization in vascular graft		
CC	surgery and is used in conjunction with angiography to administer the		
CC	angiogenic tRNA synthetase polypeptides or polynucleotides directly to		
CC	the lumen and wall of the blood vessel		
XX	Sequence 437 AA;		
XX	Query Match	100.0%;	Score 2116; DB 4; Length 437;
XX	Best Local Similarity	100.0%;	Pred. No. 1.2e-209;
XX	Matches 401; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	SNHGPDTEAEEDFVDPWTVTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQPHH	60
Db	61	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	120
Db	84	FLRRGIFFSHRDMMQVLDAYENKKEPFYLTGRGSSSEAMHVCHLIPFTKWLQDVNPV	143
QY	121	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	180
Db	144	LVIQMTDDEKYLKDLTLDQAYGDAVENAKIIACGFDINKTFIPSDLDYMGMSSEYKN	203
QY	181	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	240
Db	204	VVKIQKVTNVQVKGIFGFTDSDCIGKISPPAIQAAPSFNSPFIQFRDRTDIOCLIPCA	263
QY	241	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	300
Db	264	IDODPYFRMTRDVAPIRGYPKPAALLHSTFEPALOGAQTMSASDPNSIFLDTAKQIKT	323
QY	301	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	360
Db	324	KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLTGEL	383
QY	361	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	401
Db	384	KKALIEVLQPLIAEHQARRKEVTDEIVKERTPKLSFDFQ	424





QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEFTPRKLSFDFQ 401  
 Db 384 KKALIEVLQPLIAHQARRKEVTDEIVKEFTPRKLSFDFQ 424

## RESULT 9

ABU72385  
 ID ABU72385 standard; protein; 437 AA.

XX AC ABU72385;

XX DT 16-JUN-2003 (first entry)

XX Human mini Tryptophanyl tRNA synthetase/His tag.

XX Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytostatic;  
 KW vulnary; Rossmann fold nucleotide binding domain; chemokine; EMAP II;  
 KW vascular endothelial cell; solid tumour; myocardial infarction; enzyme;  
 KW endothelial monocyte-activating polypeptide II; tumour metastasis;  
 KW wound healing; dermal ulcer; endothelialisation; vascular graft surgery;  
 KW abdominal wound; coronary bypass surgery; gene therapy.

XX OS Homo sapiens.  
 OS Synthetic.

XX US2002182666-A1.

PD 05-DEC-2002.

XX 21-MAR-2001; 2001US-00813718.

XX 21-MAR-2001; 2001US-00813718.

XX (SCHL/) SCHIMMEL P.

XX (WAKA/) WAKASUGI K.

XX Schimmel P, Wakasugi K;

XX WPI; 2003-340974/32.

XX X-PSDB; ACA64106.

IT New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 IT Rossmann fold nucleotide binding domain or having chemokine activity  
 IT useful for e.g. for regulating angiogenesis and for treating myocardial  
 IT infarction.

IS Example 1; Page 46-47; 91pp; English.

XX The invention relates to an isolated polypeptide comprising a truncated  
 CC tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rossmann  
 CC fold nucleotide binding domain or having chemokine activity. The isolated  
 CC polypeptide is capable of regulating vascular endothelial cell function.  
 CC TyRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 CC -activating polypeptide II, a proinflammatory cytokine)-like domain and  
 CC is similar in sequence to TrpRS. Also included are a polynucleotide  
 CC encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 CC 5' and 3' deletions of the TrpRS polynucleotide, a recombinant vector  
 CC comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 CC containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 CC TrpRS antibody, producing truncations of TrpRS by treating it with a  
 CC protease, and suppressing angiogenesis, solid tumours or a condition that  
 CC would benefit from decreased angiogenesis in a mammal by administering to  
 CC the mammal the composition comprising the TrpRS truncated protein. The  
 CC isolated polypeptide is useful for the preparation of a pharmaceutical  
 CC composition for transdermal, transmucosal, enteral or parenteral  
 CC administration. The truncated tRNA synthetase polypeptide is useful for  
 CC research, diagnostic, prognostic and therapeutic applications. The tRNA  
 CC synthetase are useful for regulating vascular endothelial cell function,  
 CC particularly for regulating angiogenesis, for treating myocardial  
 CC infarction and solid tumour, and for suppressing tumour metastasis.  
 CC Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 CC agents or for treating full thickness wounds such as dermal ulcers, in  
 CC the promotion of endothelialisation in vascular graft surgery, in the

CC treatment of abdominal wounds where there is a high risk infection, in  
 CC conjunction with coronary bypass surgery by stimulating the growth of the  
 CC transplanted tissue, and its gene therapy. The present sequence is a Trp  
 CC tRNA synthetase protein (full length, truncated or mutant) with a His  
 CC affinity tag

XX SQ Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 6; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.2e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQRHH 60  
 Db 24 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQRHH 83  
 QY 61 FLERGIFFSHRDNNQVLDAYENKPFVLYTGRGSSSEAMHVGHLIPFTKWLQOVFNVP 120  
 Db 84 FLERGIFFSHRDNNQVLDAYENKPFVLYTGRGSSSEAMHVGHLIPFTKWLQOVFNVP 143  
 QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
 QY 181 VVKIQHVTNQNKGIFGFTDSDCIGKISFPAIQAAFPSPNSFPQIFRDRDIOCLIPCA 240  
 Db 204 VVKIQHVTNQNKGIFGFTDSDCIGKISFPAIQAAFPSPNSFPQIFRDRDIOCLIPCA 263  
 QY 241 IDQDPYFRMTREVPAPRIGYKPKALLHSTFFPAQGAQTKWSASDPNSIFLDTAKQIKT 300  
 Db 264 IDQDPYFRMTREVPAPRIGYKPKALLHSTFFPAQGAQTKWSASDPNSIFLDTAKQIKT 323  
 QY 301 KVNKHAFFSGGRDTEEHQFGNGCDVDVSMYLTFFLEDDKLEQIRKDYTSGLMTGEL 360  
 Db 324 KVNKHAFFSGGRDTEEHQFGNGCDVDVSMYLTFFLEDDKLEQIRKDYTSGLMTGEL 383  
 QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEFTPRKLSFDFQ 401  
 Db 384 KKALIEVLQPLIAHQARRKEVTDEIVKEFTPRKLSFDFQ 424

## RESULT 10

AAB47615

ID AAB47615 standard; protein; 484 AA.

XX AC AAB47615;

XX DT 07-JAN-2002 (first entry)

XX Human full-length TrpRS.

XX Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;  
 KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

XX Homo sapiens.

XX WO200174841-A1.

XX 11-OCT-2001.

XX 21-MAR-2001; 2001WC-US008966.

XX 31-MAR-2003; 2000US-0193471P.

XX (SCRI) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K;

XX WPI; 2001-626377/72.



Db 191 LVQMTEDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFYKN 250

QY 181 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPAIQAPSFNSPFIQPRDRDTIOCLIPCA 240

Db 251 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPAIQAPSFNSPFIQPRDRDTIOCLIPCA 310

QY 241 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 300

Db 311 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 370

QY 301 KVNKHAFSGGRDTIEHRQFGNCDVDVSMWLTFFLEDDDKLEQIRKDYTSGLMLTGE 360

Db 371 KVNKHAFSGGRDTIEHRQFGNCDVDVSMWLTFFLEDDDKLEQIRKDYTSGLMLTGE 430

QY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ 401

Db 431 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ 471

RESULT 12

AG79546

II: AAG79546 standard; protein; 484 AA.

AAG79546;

10-DEC-2002 (first entry)

Full length human TrpRS.

T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation; neovascular eye disease; age-related macular degeneration; ocular complication; diabetes; rubrotoxic glaucoma; retinopathy; prematurity; keratitis; ischaemic retinopathy; sickle cell; pathological myopia; ocular histoplasmosis; pterygia; punitate innerchoriopathy; retinal degeneration; growth factor; vascularisation; vascular endothelial cell function; angiogenesis.

Homo sapiens.

Key

Location/Qualifiers

49..484

/label= mini TrpRS

/note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"

Protein

71..484

/label= T1

/note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"

Protein

94..484

/label= T2

/note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"

Peptide

170..173

/label= Signature sequence

Misc-difference 205..207

/note= "May be replaced by Glu-Leu-Arg"

Peptide

349..353

/label= Signature sequence

WO200267970-A1.

06-SEP-2002.

22-FEB-2002; 2002WO-US005185.

23-FEB-2001; 2001US-0270951P.

(SCRI ) SCRIPPS RES INST.

Schimmel P, Wakasugi K, Friedlander M;

WPI; 2002-698635/75.

New polypeptides derived from human tryptophanyl-tRNA synthase, useful for inhibiting ocular neovascularization in a patient, or for treating neovascular eye diseases, e.g. rubrotoxic glaucoma, retinopathy, keratitis, or pterygia.

Example 1; Fig 1; 83pp; English.

This sequence represents full length human tryptophanyl-tRNA synthase (TrpRS). A cleavage product, T2, of recombinant human TrpRS is water-soluble and comprises residues 94-471 of full length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting ocular neovascularisation in a patient. This polypeptide is useful for treating neovascular eye diseases, e.g. age-related macular degeneration, ocular complications of diabetes, rubrotoxic glaucoma, retinopathy of prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell), pathological myopia, ocular histoplasmosis, pterygia, or punitate innerchoriopathy. This polypeptide is particularly useful for treating retinal degeneration to prevent the damaging effects of trophic and growth factors, and for promoting vascularisation to retard retinal degeneration by enhancing blood flow to cells. These are also useful for regulating vascular endothelial cell function, and in particular, for inhibiting angiogenesis

Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 5; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.4e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNKGEDATEAEEDFVDPWTQTSAGIDVDKLVIRFGSSKIDKELINRIERATGQRP 60

Db 71 SNKGEDATEAEEDFVDPWTQTSAGIDVDKLVIRFGSSKIDKELINRIERATGQRP 130

QY 61 FLRRGIFFSHRQMNQVLDAYENKKPFVLTGRGPSSSEAMVGHLPFTKWLQDVNVP 120

Db 131 FLRRGIFFSHRQMNQVLDAYENKKPFVLTGRGPSSSEAMVGHLPFTKWLQDVNVP 190

QY 121 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFYKN 180

Db 191 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFYKN 250

QY 181 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPAIQAPSFNSPFIQPRDRDTIOCLIPCA 240

Db 251 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPAIQAPSFNSPFIQPRDRDTIOCLIPCA 310

QY 241 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 300

Db 311 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 370

QY 301 KVNKHAFSGGRDTIEHRQFGNCDVDVSMWLTFFLEDDDKLEQIRKDYTSGLMLTGE 360

Db 371 KVNKHAFSGGRDTIEHRQFGNCDVDVSMWLTFFLEDDDKLEQIRKDYTSGLMLTGE 430

QY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ 401

Db 431 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ 471

RESULT 13

ABU72384

ID ABU72384 standard; protein; 484 AA.

AC ABU72384;

XX ABU72384;

XX 16-JUN-2003 (first entry)

XX Human full length Tryptophanyl tRNA synthetase/His tag.

XX Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytotaxic; vulnary; Rossmann fold nucleotide binding domain; chemokine; ENAP II; vascular endothelial cell; solid tumour; myocardial infarction; enzyme; endothelial monocyte-activating polypeptide II; tumour metastasis; wound healing; dermal ulcer; endothelialisation; vascular graft surgery; abdominal wound; coronary bypass surgery; gene therapy.

XX OS Homo sapiens.  
 OS Synthetic.  
 XX PF US2002182666-A1.  
 XX PF 05-DBC-2C02.  
 XX PF 21-MAR-2001; 2001US-00813718.  
 XX PF 21-MAR-2001; 2001US-00813718.  
 XX PF (SCHL/) SCHIMMEL P.  
 XX PF (WAKA/) WAKASUGI K.  
 XX PF Schimmel P, Wakasugi K;  
 XX PF WPI; 2003-340974/32.  
 XX PF N-PSDB; ACA64105.  
 XX PF New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 XX PF Rosemann fold nucleotide binding domain or having chemokine activity  
 XX PF useful for e.g. for regulating angiogenesis and for treating myocardial  
 XX PF infarction.  
 XX PF Example 1; Page 41-42; 91pp; English.  
 XX PF The invention relates to an isolated polypeptide comprising a truncated  
 XX PF tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rosemann  
 XX PF fold nucleotide binding domain or having chemokine activity. The isolated  
 XX PF polypeptide is capable of regulating vascular endothelial cell function.  
 XX PF TrpRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 XX PF activating polypeptide II, a proinflammatory cytokine)-like domain and  
 XX PF is similar in sequence to TrpRS. Also included are a polynucleotide  
 XX PF encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 XX PF 5' and 3' deletions of the TrpRS polynucleotide, a recombinant vector  
 XX PF comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 XX PF containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 XX PF TrpRS antibody, producing truncations of TrpRS by treating it with a  
 XX PF protease, and suppressing angiogenesis, solid tumours or a condition that  
 XX PF would benefit from decreased angiogenesis in a mammal by administering to  
 XX PF the mammal the composition comprising the TrpRS truncated protein. The  
 XX PF isolated polypeptide is useful for the preparation of a pharmaceutical  
 XX PF composition for transdermal, transmucosal, enteral or parenteral  
 XX PF administration. The truncated tRNA synthetase polypeptide is useful for  
 XX PF research, diagnostic, prognostic and therapeutic applications. The tRNA  
 XX PF synthetase are useful for regulating vascular endothelial cell function,  
 XX PF particularly for regulating angiogenesis, for treating myocardial  
 XX PF infarction and solid tumour, and for suppressing tumour metastasis.  
 XX PF Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 XX PF agents or for treating full thickness wounds such as dermal ulcers, in  
 XX PF the promotion of endothelialisation in vascular graft surgery, in the  
 XX PF treatment of abdominal wounds where there is a high risk infection, in  
 XX PF conjunction with coronary bypass surgery by stimulating the growth of the  
 XX PF transplanted tissue, and in gene therapy. The present sequence is a Trp  
 XX PF tRNA synthetase protein (full length, truncated or mutant) with a His  
 XX PF affinity tag  
 XX PF Sequence 484 AA;  
 XX PF Query Match 100.0%; Score 2116; DB 6; Length 484;  
 XX PF Best Local Similarity 100.0%; Pred. No. 1.4e-209;  
 XX PF Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX PF 1 SNHGPDTEAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINIRBRTGQRPHH 60  
 XX PF 71 SNHGPDTEAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINIRBRTGQRPHH 130  
 XX PF 61 FLRGGIFSHRDMMQVLDAVENKKPFYVYTGSGSSAMHVGHLIPFTKWLQDFNVVP 120  
 XX PF 131 FLRGGIFSHRDMMQVLDAVENKKPFYVYTGSGSSAMHVGHLIPFTKWLQDFNVVP 190  
 XX PF 121 LVIQWTDDEKYLWKDLTLDQAYGDAVENAKDIACGPDINKTFISDLDMGSSGYKN 180

Db 191 LVIQWTDDEKYLWKDLTLDQAYGDAVENAKDIACGPDINKTFISDLDMGSSGYKN 250  
 QY 181 VKIQKHVTENQVKGIFGFTDSDICIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240  
 Db 251 VKIQKHVTENQVKGIFGFTDSDICIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 310  
 QY 241 IDGQPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAOTKMSASDPNSSIFLTDPAKQIKT 300  
 Db 311 IDGQPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAOTKMSASDPNSSIFLTDPAKQIKT 370  
 QY 301 KVNKHAFFSGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMIFGEL 360  
 Db 371 KVNKHAFFSGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMIFGEL 430  
 QY 361 KKALIEVLQPLIAEHQARKEVTDIVKSPMTPRKLSFDFQ 401  
 Db 431 KKALIEVLQPLIAEHQARKEVTDIVKSPMTPRKLSFDFQ 471

## RESULT 14

ADE25762

ID ADE25762 standard; protein; 471 AA.

XX AC ADE25762;

XX DT 29-JAN-2004 (first entry)

XX DE Human protein differentially expressed in foam cells #39.

XX KW Human; differential expression; foam cell; LPS; lipopolysaccharide;  
 XX KW cardiovascular disease; atherosclerosis.

XX OS Homo sapiens.

XX PN US2003194721-A1.

XX PD 16-OCT-2003.

XX PF 18-SEP-2002; 2002US-00247671.

XX PR 19-SEP-2001; 2001US-0323784P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Mikita T, Shiffman D, Porter JG, Kaser VR;

XX DR WPI; 2003-875398/81.

XX DR N-PSDB; ADE25684.

XX CC Combination containing several polynucleotide that are differentially  
 XX CC expressed in foam cells and complements of the polynucleotides, useful  
 XX CC for diagnosing cardiovascular disease or atherosclerosis.

XX PS Disclosure; SEQ ID NO 166; 37pp; English.

XX CC The invention relates to a combination comprising several polynucleotides  
 XX CC having any one of 127 sequences (SI) such as the sequence of human  
 XX CC calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4  
 XX CC hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit  
 XX CC mRNA, etc., and their complements. The cDNAs are differentially expressed  
 XX CC in LPS (lipopolysaccharide)-treated foam cells. Also included are  
 XX CC obtaining an extended or full length gene from a library of nucleic acid  
 XX CC sequences, an expression vector containing the nucleic acids, a host cell  
 XX CC containing the vector, a purified polypeptide appearing as ADE25750 and  
 XX CC ADE25751, producing a protein by culturing the host cell, and a  
 XX CC composition comprising a purified antibody that specifically binds to the  
 XX CC proteins. The foam cell-expressed nucleic acids are useful for a high  
 XX CC throughput detection of differential expression of one or more  
 XX CC polynucleotides in a sample. The sample is from a subject with  
 XX CC atherosclerosis and comparison with a standard defines early, mid or late  
 XX CC stages of the disorder. The foam cell-expressed nucleic acids are useful  
 XX CC for high throughput screening of a library of molecules or compounds to

identify a ligand which binds a polynucleotide. The library is chosen from DNA molecules, peptides, proteins and RNA molecules. The protein is useful for a high throughput screening of library of molecules or compounds to identify at least one ligand which specifically binds a protein, for purifying a ligand from a sample for making an antibody. The foam cell-expressed nucleic acids are useful for diagnosing cardiovascular disorder. The foam cell-expressed nucleic acids are useful as elements on a microarray which can be used for detecting related polynucleotide in a sample, diagnosing cardiovascular disease, atherosclerosis. The present sequence represents a protein differentially expressed in JES treated foam cells.

Sequence 471 AA;

Query Match 99.3%; Score 2101; DB 7; Length 471;  
Best Local Similarity 99.5%; Pred. No. 4.7e-208;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
DB 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
QY 61 FLRRGIFFSHRDMNOVLDAENKPPYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 120  
DB 131 FLRRGIFFSHRDMNOVLDAENKPPYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 190  
QY 121 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 180  
DB 191 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 250  
QY 181 VVKIQKHVTENQVKGIFGFTSDCIGKISFPALHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 240  
DB 251 VVKIQKHVTENQVKGIFGFTSDCIGKISFPALHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 310  
QY 241 IDQDPYPRMTRDVAAPRIGYKPKALLHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 300  
DB 311 IDQDPYPRMTRDVAAPRIGYKPKALLHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 370  
QY 301 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
DB 371 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
QY 361 KKALIEVLQPLIAEHQARKEVTDEIVKEFMTPRKLSFDFQ 401  
DB 431 KKALIEVLQPLIAEHQARKEVTDEIVKEFMTPRKLSFDFQ 471

RESULT 15

AD76998  
ID ADE76998 standard; protein; 471 AA.

XX AC ADE76998;

XX 29-JAN-2004 (first entry)

DE Human protein expressed in a liver disorder #46.

XX human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;

XX tumour; liver; inflammatory disorder; immune response disorder;

XX high-throughput screening; differential gene expression; gene therapy.

CS Homo sapiens.

XX US2003108871-A1.

XX 12-JUN-2003.

XX 30-JUL-2001; 2001US-00919039.

XX 28-JUL-2000; 2000US-0222113P.

XX (KASE/J) KASER M R.

Kaser MR;

WPI; 2004-031227/03.  
N-PSDB; ADE76997.

Composition comprising several cDNAs that are differentially expressed in treated human C3A liver cell cultures, useful for treating liver disorders.

Claim 1; SEQ ID NO 163; 41pp; English.

The invention relates to a composition comprising several cDNAs that are differentially expressed in a liver disorder. The composition is useful for treating liver disorder such as hyperlipidaemia, hypertension, type II diabetes, tumours of the liver and disorders of the inflammatory and immune response. The composition is useful for a high-throughput method of screening several molecules or compounds to identify a ligand which specifically binds a cDNA. A protein encoded by the cDNA is useful for a high-throughput method for using a protein to screen several molecules or compounds to identify at least one ligand which specifically binds the protein which involves combining the protein encoded by the cDNA with several of molecules or compounds under conditions to allow specific binding, and detecting specific binding between the protein and a molecule or compound, therefore identifying a ligand which specifically binds the protein. The composition is useful for detecting and quantifying differential gene expression, can be used in gene therapy, to formulate prognosis and to design a treatment regimen and to monitor the efficacy of treatment. The present sequence represents the amino acid sequence of a protein encoded by a cDNA differentially expressed in a liver disorder.

Sequence 471 AA;

Query Match 99.3%; Score 2101; DB 8; Length 471;  
Best Local Similarity 99.5%; Pred. No. 4.7e-208;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
DB 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
QY 61 FLRRGIFFSHRDMNOVLDAENKPPYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 120  
DB 131 FLRRGIFFSHRDMNOVLDAENKPPYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 190  
QY 121 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 180  
DB 191 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 250  
QY 181 VVKIQKHVTENQVKGIFGFTSDCIGKISFPALHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 240  
DB 251 VVKIQKHVTENQVKGIFGFTSDCIGKISFPALHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 310  
QY 241 IDQDPYPRMTRDVAAPRIGYKPKALLHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 300  
DB 311 IDQDPYPRMTRDVAAPRIGYKPKALLHSTFFPALQGAQTKMSASDENSSIFLDTAKQIKT 370  
QY 301 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
DB 371 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
QY 361 KKALIEVLQPLIAEHQARKEVTDEIVKEFMTPRKLSFDFQ 401  
DB 431 KKALIEVLQPLIAEHQARKEVTDEIVKEFMTPRKLSFDFQ 471

Search completed: August 24, 2004, 17:50:32  
Job time : 130 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:01:04 ; Search time 125 Seconds  
(without alignments)  
1008.130 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPWTV.....VTDEIVKEFTPRKLSFDFO 401

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1295152 seqs, 31455058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

R result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	401	12	US-10-080-839-13
2	2116	100.0	415	9	US-09-813-718-14
3	2116	100.0	415	12	US-10-080-839-5
4	2116	100.0	415	15	US-10-240-532-14
5	2116	100.0	415	16	US-10-240-527A-14
6	2116	100.0	437	9	US-09-813-718-12
7	2116	100.0	437	12	US-10-080-839-3
8	2116	100.0	437	15	US-10-240-532-12
9	2116	100.0	437	16	US-10-240-527A-12
10	2116	100.0	471	14	US-10-126-467B-2
11	2116	100.0	471	15	US-10-295-027-1234
12	2116	100.0	484	9	US-09-813-718-10
13	2116	100.0	484	12	US-10-080-839-1
14	2116	100.0	484	15	US-10-240-532-10
15	2116	100.0	484	16	US-10-240-527A-10

16	2101	99.3	471	10	US-09-919-039-163	Sequence 163, App
17	2101	99.3	471	14	US-10-247-1671-166	Sequence 166, App
18	2101	99.3	471	16	US-10-408-765A-1235	Sequence 1235, App
19	2101	99.3	475	9	US-09-925-302-558	Sequence 558, App
20	2101	99.3	475	12	US-09-925-302-558	Sequence 558, App
21	1988	94.0	378	12	US-10-080-839-12	Sequence 12, App
22	1988	94.0	392	9	US-09-813-718-16	Sequence 16, App
23	1988	94.0	392	12	US-10-080-839-7	Sequence 7, Appl
24	1988	94.0	392	15	US-10-240-532-16	Sequence 16, Appl
25	1988	94.0	392	16	US-10-240-527A-16	Sequence 16, Appl
26	1938	91.6	475	14	US-10-205-219-65	Sequence 65, Appl
27	1326.5	62.7	408	16	US-10-437-963-120838	Sequence 120838,
28	1326	62.7	410	12	US-10-424-599-146807	Sequence 146807,
29	1302	61.5	429	12	US-10-425-114-58867	Sequence 58867, A
30	1200.5	56.7	424	14	US-10-032-585-7632	Sequence 7632, A
31	1125	53.2	433	14	US-10-128-714-8545	Sequence 8545, App
32	1107.5	52.3	456	15	US-10-320-797-3204	Sequence 3204, App
33	831	39.3	173	9	US-09-925-302-855	Sequence 855, App
34	831	39.3	173	12	US-09-925-302-855	Sequence 855, App
35	704.5	33.3	324	16	US-10-437-963-108176	Sequence 108176,
36	414.5	19.6	179	14	US-10-128-714-3545	Sequence 3545, App
37	398	18.8	137	16	US-10-437-963-145795	Sequence 145795,
38	328	15.5	85	9	US-09-813-718-45	Sequence 45, Appl
39	328	15.5	85	15	US-10-240-532-45	Sequence 45, Appl
40	328	15.5	85	16	US-10-240-527A-45	Sequence 45, Appl
41	314	14.8	113	12	US-10-424-599-146805	Sequence 146805,
42	292	13.8	85	9	US-09-813-718-46	Sequence 46, Appl
43	292	13.8	85	15	US-10-240-532-46	Sequence 46, Appl
44	292	13.8	85	16	US-10-240-527A-46	Sequence 46, Appl
45	273.5	12.9	85	9	US-09-813-718-48	Sequence 48, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-080-839-13  
; Sequence 13, Application US/10080839  
; Publication No. US20030017567A1  
; GENERAL INFORMATION:  
; APPLICANT: Wakasugi, Paul  
; APPLICANT: Friedlander, Martin  
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Arginogenesis  
; FILE REFERENCE: TSRI-813.1  
; CURRENT APPLICATION NUMBER: US/10/080.839  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-080-839-13

Query Match	100.0%	Score 2116;	DB 12;	Length 401;
Best Local Similarity	100.0%	Pred. No. 3.5e-205;		
Matches 401;	Conservative 0;	Mis-matches 0;	Indels 0;	Gaps 0;
Qy	1	SNHGPDATAEEDFVDPWTVQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPFH 60		
Db	1	SNHGPDATAEEDFVDPWTVQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPFH 60		
Qy	61	FLRGIFFSHRDMNQVLDAYENKPPFYLYTGRGSSSAMHVGHLPFTFTKWLQDVFNVP 120		
Db	61	FLRGIFFSHRDMNQVLDAYENKPPFYLYTGRGSSSAMHVGHLPFTFTKWLQDVFNVP 120		
Qy	121	LVIOMTDEKYLKDLTLDOAYGDAVENAKDIACGPDINKTFISDLDMGSSGFYKN 180		
Db	121	LVIOMTDEKYLKDLTLDOAYGDAVENAKDIACGPDINKTFISDLDMGSSGFYKN 180		

```

181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 240
181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 240
241 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 300
241 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 300
301 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
301 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
361 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 401
361 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 401

```

## RESULT 2

```

US-09-813-718-14
Sequence 14, Application US/09813718
Publication No. US20020182666A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813,718
NUMBER OF SEQ ID NOS: 58
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 14
LENGTH: 415
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: human
OTHER INFORMATION: supermini Trprs in pET20B

```

## US-09-813-718-14

```

Query Match 100.0%; Score 2116; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60
2 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61
61 FLRRGIFFSHRDMNQVLDAYENKPKPYLYTGRGPSSEAMHVGHLPFTTKWLQDVFNVP 120
62 FLRRGIFFSHRDMNQVLDAYENKPKPYLYTGRGPSSEAMHVGHLPFTTKWLQDVFNVP 121
121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
122 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181
181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 240
182 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 241
241 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 300
242 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 301
301 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
302 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361
361 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 401
362 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 402

```

## RESULT 3

```

US-10-080-839-5
Sequence 5, Application US/10080339
Publication No. US20030017564A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
APPLICANT: Friedlander, Martin
TITLE OF INVENTION: Tryptophanyl-tRNA Synthetase Derived
FILE REFERENCE: TSRI-813.1
CURRENT APPLICATION NUMBER: US/10/080,839
CURRENT FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: 60/270,951
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 415
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cleavage Product T1 of recombinant human Trprs
US-10-080-839-5

```

```

Query Match 100.0%; Score 2116; DB 12; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60
DB 2 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61
QY 61 FLRRGIFFSHRDMNQVLDAYENKPKPYLYTGRGPSSEAMHVGHLPFTTKWLQDVFNVP 120
DB 62 FLRRGIFFSHRDMNQVLDAYENKPKPYLYTGRGPSSEAMHVGHLPFTTKWLQDVFNVP 121
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 122 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181
QY 181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 240
DB 182 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDTIQCLIPCA 241
QY 241 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 300
DB 242 IDODPYFMTDRVAPRIGYKPKALLHSTFPALQAQTKMSASDPSNSSFILDTAKQIKT 301
QY 301 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
DB 302 KVNKHAFGSGRDTIEHRQFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361
QY 361 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 401
DB 362 KKALIEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFDQ 402

```

## RESULT 4

```

US-10-240-532-14
Sequence 14, Application US/10240532
Publication No. US20040009163A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: TSRI 720.1
CURRENT APPLICATION NUMBER: US/10/240,532
CURRENT FILING DATE: 2002-03-30
PRIOR APPLICATION NUMBER: PCT/US01/08975
PRIOR APPLICATION NUMBER: 2001-03-21

```

;; PRIOR APPLICATION NUMBER: US 60/193,471  
;; PRIOR APPLICATION NUMBER: 2000-03-31  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 14  
;; LENGTH: 415  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: human  
;; OTHER INFORMATION: supermini TrpRS in pET20B  
US-10-240-527A-14

Query Match  
Best Local Similarity 100.0%; Score 2116; DB 15; Length 415;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 61  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 62 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 121  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 180  
ID 122 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 181  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDTDIQCLIPCA 240  
ID 182 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDTDIQCLIPCA 241  
CY 241 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 300  
ID 242 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 301  
CY 301 KVNKHAFFSGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
ID 362 KVNKHAFFSGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
CY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSDFDQ 401  
ID 362 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSDFDQ 402

RESULT 5  
US-10-240-527A-14  
;; Sequence 14, Application US/10240527A  
;; Publication No. US20040152079A1  
;; GENERAL INFORMATION:  
;; APPLICANT: SCHIMMEL, Paul  
;; APPLICANT: WAKASUGI, Keisuke  
;; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase  
;; TITLE OF INVENTION: Polypeptides Useful for the Regulation of Angiogenesis  
;; FILE REFERENCE: TSRI 720.2  
;; CURRENT APPLICATION NUMBER: US/10/240,527A  
;; CURRENT FILING DATE: 2002-09-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/08966  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/193,471  
;; PRIOR FILING DATE: 2000-03-31  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 14  
;; LENGTH: 415  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Human supermini TrpRS in pET20B  
US-10-240-527A-14

Query Match  
100.0%; Score 2116; DB 16; Length 415;

Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 61  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 62 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 121  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 180  
ID 122 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 181  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDTDIQCLIPCA 240  
ID 182 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDTDIQCLIPCA 241  
CY 241 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 300  
ID 242 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 301  
CY 301 KVNKHAFFSGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
ID 362 KVNKHAFFSGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
CY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSDFDQ 401  
ID 362 KKALIEVLOPLIAEHQARRKEVTDEIVKFMTPRKLSDFDQ 402

## RESULT 6

US-09-813-718-12  
;; Sequence 12, Application US/09813718  
;; Publication No. US20020182666A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Schimmel, Paul  
;; APPLICANT: Wakasugi, Keisuke  
;; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
;; TITLE OF INVENTION: The Regulation of Angiogenesis  
;; FILE REFERENCE: 00-221  
;; CURRENT APPLICATION NUMBER: US/09/813,718  
;; CURRENT FILING DATE: 2001-03-21  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 12  
;; LENGTH: 437  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: human mini  
;; OTHER INFORMATION: TrpRS in pET20B  
US-09-813-718-12

Query Match  
Best Local Similarity 100.0%; Score 2116; DB 9; Length 437;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 83  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 84 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 143  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 180  
ID 144 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLYMGSSGFYKN 203  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDTDIQCLIPCA 240

Db 204 VKIQKHVTNQGKIFGFTSDCIGKISFPALQAAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424  
RESULT 7  
US-10-080-839-3  
; Sequence 3, Application US/10080839  
; Publication No. US20030017564A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; APPLICANT: Friedlander, Martin  
; TITLE OF INVENTION: tryptophanyl-tRNA Synthetase Derived  
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Angiogenesis  
; FILE REFERENCE: TSRI-813.1  
; CURRENT APPLICATION NUMBER: US/10/080,839  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/270,951  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: human mini TrpS in pET20B  
US-10-080-839-3

Query Match 100.0%; Score 2116; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
Db 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
QY 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHILIPFTKWLQDVNVP 120  
Db 84 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHILIPFTKWLQDVNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIQKHVTNQGKIFGFTSDCIGKISFPALQAAPSFNSFPQIPRDRDIOCLIPCA 240  
Db 204 VKIQKHVTNQGKIFGFTSDCIGKISFPALQAAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424  
RESULT 8  
US-10-240-532-12  
; Sequence 12, Application US/10240532  
; Publication No. US20040009163A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
; TITLE OF INVENTION: The Regulation of Angiogenesis  
; FILE REFERENCE: TSRI 720.1  
; CURRENT APPLICATION NUMBER: US/10/240,532  
; PRIOR FILING DATE: 2002-03-30  
; PRIOR APPLICATION NUMBER: PCT/US01/08975  
; PRIOR APPLICATION NUMBER: 2001-03-21  
; PRIOR APPLICATION NUMBER: JS 60/193,471  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human mini  
; OTHER INFORMATION: TrpS in pET20B  
US-10-240-532-12

Query Match 100.0%; Score 2116; DB 15; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
Db 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
QY 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHILIPFTKWLQDVNVP 120  
Db 84 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHILIPFTKWLQDVNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIQKHVTNQGKIFGFTSDCIGKISFPALQAAPSFNSFPQIPRDRDIOCLIPCA 240  
Db 204 VKIQKHVTNQGKIFGFTSDCIGKISFPALQAAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAQTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTIEHRQFGGNCDDVDSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

RESULT 9  
US-10-240-527A-12  
; Sequence 12, Application US/10240527A  
; Publication No. US20040152079A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase  
; TITLE OF INVENTION: Polypeptides Useful for the Regulation of Angiogenesis  
; FILE REFERENCE: TSRI 720.2  
; CURRENT APPLICATION NUMBER: US/10/240,527A  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: PCT/US01/08966

; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/193,471  
 ; PRIOR FILING DATE: 2003-03-31  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Human mini TrpRS in pET20B  
 U3-10-240-527A-12

Query Match 100.0%; Score 2116; DB 16; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 4e-205;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q'	1	SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP	60
D	24	SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP	83
Qy	61	FLRGIFFSHRDMNQVLDAVENKPFYLYTGRGSPSEAMHVGHLIPFFITKWLQDVFNVP	120
D	84	FLRGIFFSHRDMNQVLDAVENKPFYLYTGRGSPSEAMHVGHLIPFFITKWLQDVFNVP	143
Qy	121	LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFFISDLDYMGSSGFYKN	180
D	144	LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFFISDLDYMGSSGFYKN	203
Qy	181	VVKIQKHVTNFQVKGIFGFTDSDCIGKISFPALQAQATKMSADPNSSIFLDTAKQIKT	240
D	204	VVKIQKHVTNFQVKGIFGFTDSDCIGKISFPALQAQATKMSADPNSSIFLDTAKQIKT	263
Qy	241	IDQPYFMTDVAIPRIGYKPKALLHSFFPALQAQATKMSADPNSSIFLDTAKQIKT	300
D	264	IDQPYFMTDVAIPRIGYKPKALLHSFFPALQAQATKMSADPNSSIFLDTAKQIKT	323
Qy	301	KVNKHAFGSGRDTIEHROFGNCDVVSFMYLTFPLEDDKLEQIRKDYTGAMLTGEL	360
D	324	KVNKHAFGSGRDTIEHROFGNCDVVSFMYLTFPLEDDKLEQIRKDYTGAMLTGEL	383
Qy	361	KKALIEVLQPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ	401
D	384	KKALIEVLQPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ	424

RESULT 10  
 US-10-126-467B-2  
 ; Sequence 2, Application US/10126467B  
 ; Publication No. US20030659797A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paley, Elena  
 ; TITLE OF INVENTION: ANIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE  
 ; FILE REFERENCE: PALL-111  
 ; CURRENT APPLICATION NUMBER: US/10/126,467B  
 ; CURRENT FILING DATE: 2002-11-19  
 ; PRIOR APPLICATION NUMBER: 60/284,980  
 ; PRIOR FILING DATE: 2001-04-19  
 ; PRIOR APPLICATION NUMBER: 09/513,895  
 ; PRIOR FILING DATE: 2000-02-28  
 ; PRIOR APPLICATION NUMBER: 09/384,869  
 ; PRIOR FILING DATE: 1999-08-27  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 471  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 U3-10-126-467B-2

Query Match 100.0%; Score 2116; DB 14; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-205;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP	60
D	71	SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP	130
Qy	61	FLRGIFFSHRDMNQVLDAVENKPFYLYTGRGSPSEAMHVGHLIPFFITKWLQDVFNVP	120
D	131	FLRGIFFSHRDMNQVLDAVENKPFYLYTGRGSPSEAMHVGHLIPFFITKWLQDVFNVP	190
Qy	121	LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFFISDLDYMGSSGFYKN	180
D	191	LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFFISDLDYMGSSGFYKN	250
Qy	181	VVKIQKHVTNFQVKGIFGFTDSDCIGKISFPALQAQATKMSADPNSSIFLDTAKQIKT	240
D	251	VVKIQKHVTNFQVKGIFGFTDSDCIGKISFPALQAQATKMSADPNSSIFLDTAKQIKT	310
Qy	241	IDQPYFMTDVAIPRIGYKPKALLHSFFPALQAQATKMSADPNSSIFLDTAKQIKT	300
D	311	IDQPYFMTDVAIPRIGYKPKALLHSFFPALQAQATKMSADPNSSIFLDTAKQIKT	370
Qy	301	KVNKHAFGSGRDTIEHROFGNCDVVSFMYLTFPLEDDKLEQIRKDYTGAMLTGEL	360
D	371	KVNKHAFGSGRDTIEHROFGNCDVVSFMYLTFPLEDDKLEQIRKDYTGAMLTGEL	430
Qy	361	KKALIEVLQPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ	401
D	431	KKALIEVLQPLIAEHQARRKEVTDEIVKFMTPRKLSFDFQ	471

RESULT 11  
 US-10-295-027-1234  
 ; Sequence 1234, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevezi, Peter A.  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Murray, Richard  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 ; FILE REFERENCE: 018501-012500US  
 ; CURRENT APPLICATION NUMBER: US/10/295,027  
 ; CURRENT FILING DATE: 2002-11-13  
 ; PRIOR APPLICATION NUMBER: US 09/663,733  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: US 60/350,666  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/335,394  
 ; PRIOR FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/332,464  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/334,393  
 ; PRIOR FILING DATE: 2001-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/340,376  
 ; PRIOR FILING DATE: 2001-12-14  
 ; PRIOR APPLICATION NUMBER: US 60/347,211  
 ; PRIOR FILING DATE: 2002-01-08  
 ; PRIOR APPLICATION NUMBER: US 60/347,349  
 ; PRIOR FILING DATE: 2002-01-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,250  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: US 60/356,714  
 ; PRIOR FILING DATE: 2002-02-13  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1386  
 ; SOFTWARE: PatentIn Ver. 2.1

```
/ SEQ ID NO 1234
/ LENGTH: 471
/ TYPE: PRT
/ ORGANISM: Homo sapiens
U.-10-295-027-1234

Query Match      100.0%; Score 2116; DB 15; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
DB 71 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 120
DB 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 190
QY 121 LV1QMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGPDINKTFIFSDLYMGSSGGFYKN 180
DB 191 LV1QMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGPDINKTFIFSDLYMGSSGGFYKN 250
QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTMSASDPNSSIFLTDTAQIKT 300
DB 311 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTMSASDPNSSIFLTDTAQIKT 370
QY 301 KVKHAFSGGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 360
DB 371 KVKHAFSGGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 430
QY 361 KXALIEVLQPLIAEHQARRKEVTDEIVKEPMTPRKLSDFDQ 401
DB 431 KXALIEVLQPLIAEHQARRKEVTDEIVKEPMTPRKLSDFDQ 471

RESULT 13
US-10-080-839-1
; Sequence 1, Application US/10080839
; Publication No. US20030017564A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakaugi, Keisuke
; APPLICANT: Friedlander, Martin
; TITLE OF INVENTION: Tryptophanyl-tRNA Synthetase Derived
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Arg-ogenesis
; FILE REFERENCE: TSRI-813.1
; CURRENT APPLICATION NUMBER: US/10/080,839
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,951
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant human trpRS
US-10-080-839-1

Query Match      100.0%; Score 2116; DB 12; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
DB 71 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 120
DB 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 190
QY 121 LV1QMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGPDINKTFIFSDLYMGSSGGFYKN 180
DB 191 LV1QMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGPDINKTFIFSDLYMGSSGGFYKN 250
QY 181 VKYIQKHVTFNQVKGIFGFTDSDICIGKISPPALQGAQTSASDPNSSIFLTDTAQIKT 240
DB 251 VKYIQKHVTFNQVKGIFGFTDSDICIGKISPPALQGAQTSASDPNSSIFLTDTAQIKT 310
QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTMSASDPNSSIFLTDTAQIKT 300
DB 311 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTMSASDPNSSIFLTDTAQIKT 370
QY 301 KVKHAFSGGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 360
DB 371 KVKHAFSGGRDTIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 430

RESULT 12
US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US20020192666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakaugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 90-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human
; OTHER INFORMATION: full-length TrpRS in pET20B
US-09-813-718-10

Query Match      100.0%; Score 2116; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
DB 71 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 120
DB 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSPSEAMVGHILPIFTKWLQDVNVP 190
```



```

Db 371 KVNKAFSGGRDTIEHRQFGNCVDVSPXYLTFFLEDDDKLEQIRKQYTSGLMTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

RESULT 14
US-10-240-532-10
; Sequence 10, Application US/10240532
; Publication No. US20040009163A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: TSRI 720.1
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/10/240,532
; PRIOR APPLICATION NUMBER: PCT/US01/08975
; PRIOR APPLICATION NUMBER: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/193,471
; PRIOR APPLICATION NUMBER: 2000-03-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: human
; OTHER INFORMATION: full-length TrpRS in pET20B
US-10-240-532-10

Query Match 100.0%; Score 2116; DB 15; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 60
IB 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 130
CY 61 FLRRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMVGHLPFFITKWLQDVNPV 120
IB 131 FLRRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMVGHLPFFITKWLQDVNPV 190
CY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
IB 191 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
CY 181 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240
IB 251 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 310
CY 241 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 300
IB 311 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 370
CY 301 KVNKAFSGGRDTIEHRQFGNCVDVSPXYLTFFLEDDDKLEQIRKQYTSGLMTGEL 360
IB 371 KVNKAFSGGRDTIEHRQFGNCVDVSPXYLTFFLEDDDKLEQIRKQYTSGLMTGEL 430
CY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

```

```

RESULT 15
US-10-240-527A-10
; Sequence 10, Application US/10240527A
; Publication No. US2004015209A1
; GENERAL INFORMATION:
; APPLICANT: SCHIMMEL, Paul

```

```

; APPLICANT: WAKASUGI, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase
; FILE REFERENCE: Polypeptides Useful for the Regulation of Angiogenesis
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/10/240,527A
; PRIOR APPLICATION NUMBER: PCT/US01/08966
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/193,471
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Versi0c 4.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Human full-length TrpRS in pET20B
US-10-240-527A-10

Query Match 100.0%; Score 2116; DB 16; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 60
DB 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 130
CY 61 FLRRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMVGHLPFFITKWLQDVNPV 120
DB 131 FLRRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMVGHLPFFITKWLQDVNPV 190
CY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 191 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
CY 181 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240
DB 251 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 310
CY 241 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 300
DB 311 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 370
CY 301 KVNKAFSGGRDTIEHRQFGNCVDVSPXYLTFFLEDDDKLEQIRKQYTSGLMTGEL 360
DB 371 KVNKAFSGGRDTIEHRQFGNCVDVSPXYLTFFLEDDDKLEQIRKQYTSGLMTGEL 430
CY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

```

Search completed: August 24, 2004, 19:21:54  
Job time : 127 secs

